Characterisation of Mucosal Associated Invariant T-cells and MR1 in Ruminants

Nicholas Graham Goldfinch

A thesis submitted for the degree of Doctor of Philosophy
at The University of Edinburgh, 2010

Doctor of Philosophy – The University of Edinburgh – 2010

Table of Contents

Index of Figures

Index of Tables

Declaration

I declare that the work presented in this thesis is the product of my own efforts, and the practical research on which it is based is my own except where stated in the text and in the acknowledgement section. This work has not been previously submitted for any other degree or professional qualification.

Nicholas G. Goldfinch

Acknowledgements

I must first express my gratitude to my principal supervisor, Professor Ivan Morrison, who has provided me with the inspiration to fulfill this project, as well as countless discussions of advice and encouragement on both my practical work and my thesis writing.

I would also like to express my gratitude to my colleagues in the lab, who have provided me with frequent advice, practical guidance and reagents, most notably Alison Burrells and Dr. Niall MacHugh, and also Dr. Jane Hart, Dr. Tim Connelley, Jie Yang, and Miray Tonk. I am also greatly indebted to my friends, Dr. Frank Katzer, Dr. Gerry McLachlan, Dr. Keith Ballingall, and Dr. Neil Anderson, for their constant encouragement, technical advice and support.

I must also express my thanks to Dr. Idliko van Rhijn, University of Utrecht, for collaborations on bovine MAIT and MR1 work, as well as Dr. Kim Willoughby, Moredun Research Institute, and Mark Lynch, Applied Biosystems, for their excellent advice on qPCR assay design and troubleshooting,

Lastly, but most importantly, I must thank my parents, Judy and Paul, and my family, Gillian, Lewis and Ben, for their love, support, advice and patience. This is for you.

Abstract

Mucosal associated invariant T-cells (MAIT) are a phylogenetically conserved subset of alpha/beta T-cells with natural killer-like (NK) activity. MAIT are defined by the expression of an invariant T-cell receptor alpha (TCRα) chain; in mice and humans this chain uses the orthologous mVα19/hVα7.2-Jα33 genes respectively. Available evidence indicates that MAIT are restricted by MR1, a highly conserved MHC class I-related molecule, and that their development is dependent on B lymphocytes. They appear to constitute part of the innate immune response, but their precise functional role is poorly understood. This study aimed to characterise MAIT and MR1 in ruminants, and to further the knowledge and understanding of these unique cells.

Using PCR primers based on partial database sequences, orthologous full-length TCRα chains were identified in circulating bovine and ovine T cells. The germline elements of the respective α chains were identified and their overall frequency of expression within the bovine TCRα repertoire determined. Experiments using the orthologous TCRα chain as a marker for MAIT cells to examine expression in bovine and ovine blood and various tissues showed that spleen and mesenteric lymph nodes contained the highest frequency of MAIT cells. Use of the same technique to study levels of this marker in cattle of different ages revealed very low numbers of MAIT cells in neonatal animals, followed by a marked increase in the first 3 weeks of life. Analyses of MAIT TCRα expression in different T cell subsets showed that, unlike mice and humans in which MAIT cells are predominantly within the $CD4^-$ /$CD8^-$ T-cell population, MAIT cells in bovine blood are predominantly $CD8^+$. Full-length cDNAs were isolated for bovine and sheep MR1 and their sequences were found to display marked cross-species conservation. Using a specific PCR, MR1 was shown to be expressed in peripheral blood and by different lineages of Theileria-transformed cells. Alternatively-spliced transcripts of MR1 were detected in both cattle and sheep and several of these retained an intact open-reading frame. Constructs of bovine MR1 and an MR1/MHC chimera were prepared in a eukaryotic expression vector but these failed to give detectable cell surface expression following transfection into Cos-7, despite positive intracellular expression.

1 General Introduction

1.1 The Innate and Adaptive Immune Systems

1.1.1 General Overview

The host immune defence consists of two major, interacting, elements – the innate (non-specific) and adaptive (specific) immune systems. These systems both consist of humoral and cellular components but play different roles in relation to defence from infection. The innate immune response provides a rapid but non-specific response to challenge, mobilizing defences, many of which are constitutively-present and prepared, against a wide variety of organisms. Physical and chemical barriers of the body also function as part of the innate immune system. In contrast, the adaptive immune response is slower but more specific, reacting only against the original insulting antigen and further providing immunological memory and protection against repeated infection with the same organism. Despite their differences there is interplay between the two systems and the innate immune system has an important role in stimulating the adaptive response to infection.

1.1.2 The Innate Immune System

The innate immune system provides an immediate response to infection. The most immediate challenge is the 'anatomical' barrier. This consists of both the physical barrier of organs and cells, such as the skin and epithelial layer, and also the mechanical barriers within organs preventing micro-organism build-up or access, such as peristalsis within the gut, bronchial ciliary movement, or fluid and mucous on mucosal surfaces. Further support to this anatomical barrier is provided by both biological factors, such as competitive commensal flora, and chemical secretions which can inhibit the growth of, and actively damage, invading micro-organisms.

If the anatomical barrier is breached, then the humoral element of the innate system initiates an acute inflammatory response at the site of infection. This is mediated by

humoral factors in serum or released at the infection-site, and results in a number of changes including oedema and recruitment and activation of phagocytic cells. One of the major components of this barrier is the complement system. This non-specific defence mechanism comprises a cascade of serum proteins, the final products of which are able to lyse bacteria, as well as opsonising bacteria for rapid phagocytosis. Complement also plays a role in recruiting phagocytic cells to the site of injury and increasing the permeability of the vascular membrane, causing oedema and allowing rapid access for these cells to the specific site of infection. The coagulation system may also be activated as part of the innate immune system's humoral barrier – this attracts phagocytic cells and increases vascular permeability. In addition, elements of the coagulation system can have direct antimicrobial effects. Other components of the humoral innate response also provide direct antimicrobial or antiviral effects – lactoferrin and transferrin limit bacterial growth, lysozymes damage bacterial cell walls, interferon reduces viral replication within cells, and interleukin-1 stimulates the acute phase protein response.

The principal part of the innate immune response is the cellular component, the major functions of which include phagocytosis of invading organisms and killing of intracellular pathogens. The major cell-types involved in this are neutrophils and macrophages, both local tissue macrophages and circulating monocytes which mature into macrophages in response to infection. Circulating cells are attracted by chemical signals from local innate inflammatory mediators and cross the vascular endothelial barrier by diapedesis. The local chemical signals then attract these cells to the specific site of infection by chemotaxis and also stimulate them to provide enhanced phagocytic activity and killing. Phagocytic cells bind to foreign or infectious agents and stimulate phagocytosis via a variety of cell-surface receptors, including complement receptors and scavenger receptors. Activated macrophages also function through killing of intracellular pathogens by induction of nitric oxide and generation of free oxygen radicals. Eosinophils can also be effector cells of the innate immune response by phagocytosing and killing some parasites. Interaction of pathogens or their products with pattern recognition receptors (PRRs) including the Toll-like receptors (TLRs), which detect pathogen-associated molecular patterns

(PAMPs), results in activation of various signalling pathways that activate the cells and enables them to promote the induction of adaptive immune responses.

Natural killer (NK) cells also play a major part in the cellular component of the innate immune response. These are a type of large granular lymphocyte which do not express either the T-cell or B-cell antigen-specific surface receptors. NK cells can undertake non-specific killing of infected or neoplastic cells. They kill cells by inducing apoptosis following release of perforin and granzymes, in a manner similar to $CD8^+$ T cells.

The innate immune system also stimulates and mediates activation of the adaptive immune system in response to infection, see section 1.1.4.

(based on publications by Janeway 1993; Medzhitov and Janeway 2000; Janeway 2001; Alberts et al. 2002; Janeway and Medzhitov 2002; Beutler 2004; Mayer 2007)

1.1.3 The Adaptive Immune System

The adaptive immune system is unique to jawed vertebrates. It allows the immune system to recognise and remember specific pathogens and therefore mount a more rapid and effective immune response to subsequent infections from the same pathogen. This specific immune system has the ability to detect millions of different antigens due to its large and extremely diverse antigen-receptor repertoire.

The main effector cells of the adaptive immune system are T and B lymphocytes, each of which comprises cells expressing a repertoire of structurally-unique antigen-receptors. If an individual lymphocyte encounters an antigen that can bind to its receptor, the cell is activated and induced to proliferate and differentiate, thereby providing a specific response to the pathogen. B-cells and T-cells are derived from the same pluripotential haematopoietic stem cell and together compose roughly 20-40% of the human leucocyte population, which continuously recirculate between the blood and the lymphatic system. These cells can either be in a naïve state, having

not yet encountered an appropriate antigen, an effector state, when actively responding to foreign antigen, or a memory state, having reverted to a resting state following previous activation (publications by Janeway 2001; Alberts et al. 2002).

B-cells are the major cells of humoral immunity, being primarily involved in the production of antibodies. B-cell surface receptors (immunoglobulins) recognise and bind to antigens in their native state which, along with co-stimulation from 'helper' T-cells, stimulates differentiation of the B-cell into an 'effector' plasma cell, producing specific antibody. Most antibody-producing cells are short-lived but a small proportion of them survive as long-lasting 'memory' cells, allowing a more rapid antibody response following re-infection with the same pathogen.

T-cells are the principal cells involved in specific cell-mediated immunity. Each T-cell possesses a surface T-cell receptor (TCR) which recognises processed antigenic peptides presented in association with major histocompatibility complex (MHC) molecules (Zinkernagel and Doherty 1974) on the surface of antigen presenting cells (APC). All nucleated cells can theoretically function as APC, although the properties of dendritic cells (DC) and B-cells enable them to be highly effective APC and they are therefore termed 'professional' APC (reviewed by Banchereau and Steinman 1998). T-cells undergo development in the thymus and give rise to several phenotypically distinct subsets of T-cells. $CD8^+$ T-cells, when activated, secrete a number of immuno-regulatory cytokines and mediate killing of virus-infected or neoplastic cells. These cells recognise antigenic proteins processed through an endogenous pathway and presented by MHC class I molecules (reviewed by Harty et al. 2000). In contrast, $CD4^+$ T-cells recognise antigenic proteins (usually taken up from the extracellular environment) processed through an exogenous pathway to generate peptides that are presented by MHC class II molecules (Fearon and Locksley 1996). Upon activation $CD4^+$ T-cells release a range of cytokines which influence the cell activity of many other cells, including the original APC and cytotoxic T-cells. The $CD4^+$ T-cell response can vary depending on the antigenic stimulus; the two extremes of function are classified as Th1 or Th2 responses and are defined by the cytokines they produce. Th1 responses are primarily induced against

intracellular pathogens (virus/bacteria/protozoa), and are characterised by the production of interferon-gamma (IFNγ) and tumour necrosis factor-alpha (TNF-α). They can stimulate signalling pathways in the target cells that result in intracellular killing of pathogens. In contrast the Th2 response is more effective against extracellular bacteria, parasites and toxins, and is characterised by production of cytokines interleukin 4 (IL-4), IL-5 and IL-13, which mediate inflammatory responses and changes at mucosal surfaces that facilitate expulsion of parasites. $CD4^+$ T cells also play an important role in providing help for generating antibody responses and different cytokines influence the isotypes of antibodies that are produced (Murray 1998; Kidd 2003).

An additional αβ T-lymphocyte subset are the 'regulatory' T-cells, which play a major role in limiting and suppressing the host immune response and controlling excessive inflammation and autoimmunity (Beissert et al. 2006).

Gamma/delta T-cells form a separate subpopulation expressing a different TCR composed of gamma and delta chains. These cells share properties of helper T-cells, cytotoxic T-cells and NK cells, and therefore function as part of both the innate and adaptive immune responses. Their TCR repertoire is more limited than αβ T-cells due to a much more limited range of δ variable-region genes (Raulet 1989; Hein and Mackay 1991).

1.1.4 Links between the innate & adaptive immune systems

The innate immune system plays a vital role in the control and activation of the adaptive immune response. Recognition of antigen via the specific surface receptor is a key event in stimulating a B or T cell response. However, in the case of T cells, binding of the antigen-receptor alone is not sufficient to stimulate a response. Co-stimulation by engagement of other surface molecules, such as CD80 and CD86, with ligands on antigen presenting cells, is also required, while the local cytokine environment can also significantly affect the nature of T-cell response induced

(reviewed by Medzhitov and Janeway 2000; Beutler 2004). The principal cell type involved in presenting antigen to T cells during primary responses is the DC, which is capable of expressing high levels of co-stimulatory ligands and MHC molecules. However, up-regulation of co-stimulatory ligands, as well as cytokines involved in triggering T cell activation, on DCs is dependent on activation of the DC via one or more cell receptors (either on the cell surface or within endosomes) that recognise PAMPs. The latter is a key component of the innate immune response and involves a number of PAMP receptors, including the TLRs, which recognise different molecular patterns, such as bacterial lipopolysaccharide (endotoxin), flagellin, or peptidoglycan, as well as double-stranded RNA, or unmethylated DNA, associated with viruses. The particular innate receptors that are activated influence which DC surface co-receptors and cytokines are up-regulated. In this way, the innate immune response also affects how the T cells differentiate and therefore the nature of the T cell response. Hence, the innate response plays a key role both in promoting the adaptive response but also ensuring that the responses generated are appropriate for the particular pathogen (reviewed by Banchereau and Steinman 1998; Reis e Sousa 2004).

NK cells, in addition to recognising infected or neoplastic cells, can also interact with DCs, and this NK-cell/DC crosstalk plays a significant role in the coordination of innate and adaptive immune responses. Activated NK cells can either induce DC maturation directly, or in association with suboptimal levels of microbial signals; this is dependent on both cell-cell contact, and local cytokine release, including TNFα and IFNγ. These direct influences on DC activation can therefore indirectly influence the T cell response (Walzer et al. 2005).

A more recently hypothesised link between the innate and adaptive immune systems has been provided by the discovery of small subsets of T-cells defined by the use, in each case, of an invariant TCR chain (Lantz and Bendelac 1994; Han et al. 1999; Allez et al. 2002). The highly conserved structure of the TCRs expressed by these subsets, and their relatively high frequency compared to individual conventional T-cell TCR specificities, may allow them to generate rapid responses and function

essentially like cells of the innate immune system, such as NK cells. The presence of NK markers on the cell surface of some of these subsets may also confer similar innate properties to these T-cells. The TCRs of these T-cell subsets also show marked conservation between mammalian species, suggesting that they play a significant evolutionarily conserved role in the host's immune defence.

1.2 The T-cell

As described in 1.1.3, T-lymphocytes are divided into subsets according to effector function. Development of these lineages occurs within the thymus during the process of TCR gene rearrangement when T cells are selected based on their affinity for class I or class II MHC proteins with bound self peptides. Hence the defining molecule on these cells is the TCR, which functions both in antigen recognition and, by associating with a number of other polypeptides (CD3), in mediating activation of the T-cell upon antigen recognition.

1.2.1 The αβ TCR

1.2.1.1 TCR Composition and Interaction with MHC

The TCR is a heterodimeric cell surface protein composed of two di-sulphide linked polypeptide chains. The majority of T-cells that recognise antigenic peptides express an α and a β chain, while a second type of TCR made up of γ and δ chains defines a separate subset of T-cells which account for <10% of peripheral human T-cells but up to 30-50% of peripheral T-cells in neonatal ruminants (Clevers et al. 1990; Hein and Mackay 1991). The precise antigenic specificity of these γδ T cells is still poorly defined. The α and β polypeptide chains have similar domain structures, each being composed of a short cytoplasmic tail, a trans-membrane region, a constant (C) domain and an N-terminal variable (V) domain. The variable domains combine to form the antigen recognition site, which is composed of six superficial loops, three contributed by each chain, that interact with the antigenic peptide and/or the presenting MHC molecule. The α and β chain transcripts contain short hypervariable segments, referred to as complementarity determining regions (CDR), which encode the antigen-binding loops. The constant and trans-membrane domains form a short 'hinge' region at their junction which is the site of one of the cysteine residues forming the inter-chain disulphide bonds; the other is within the variable domain. At the cell surface the entire αβ TCR is non-covalently linked via its trans-membrane

domain to the invariant accessory chains of the CD3 complex which function in signal transduction following ligation of the TCR (Allison and Lanier 1987; Chothia et al. 1988; Davis and Bjorkman 1988).

T-cell receptors are encoded by four distinct gene families: α, β, γ, δ, contained within three chromosomal loci: α/δ, β, γ. The TCR genes consist of separate germline segments that are somatically rearranged and joined during T-cell differentiation to encode the variable and constant domains of the TCR. Variable region genes are assembled from three distinct germline segments - variable (V), diversity (D) and joining (J) – except for the α chain, which uses V and J only. The CDR loops, which interact with the MHC molecule and the antigen it presents, are encoded either within the V gene itself (CDR1 and CDR2) or by the conjunction of V (D) and J (CDR3) which allows maximal utilisation of the sequence polymorphism within the TCR chains (Allison and Lanier 1987; Davis and Bjorkman 1988; Glusman et al. 2001; Mackelprang et al. 2006).

The CDR3 regions from either chain are sited at the centre of the TCR combining site, flanked by their respective CDR1/2 regions. The inter-CDR3 region can either be flat or contain a hollow which can entirely engulf side-chains of MHC-presented peptides (Garcia et al. 1998). The overall combining site surface of the TCR is relatively flat and this is consistent with the flat/undulating surface of the MHC (Garcia et al. 1996; reviewed by Garcia et al. 1999). MHC/peptide complexes present a complex surface of peptide and MHC atoms to the TCR 'combining site', in which only about one-third of the bound peptide surface is exposed, and the TCR interacts with this whole complex of MHC and peptide. The TCR is orientated diagonally across the MHC surface, which accords a relatively constant orientation of the Vα and Vβ domains and CDR regions (Garboczi et al. 1996; reviewed by Bjorkman 1997; reviewed by Garcia and Adams 2005). The CDR3 loops of both chains interact with the MHC-bound peptide – the Vα CDR3 interacts with a central position of the peptide, while the Vβ CDR3 lies over the peptide's C-terminal region either as a flat surface or engulfing the peptide in a central hollow (Garboczi et al. 1996; Bjorkman 1997; Hennecke and Wiley 2001). The CDR2 regions lie directly

on top of the MHC α2 (Vα CDR2) and α1 (Vβ CDR2) helices and do not contact the peptide. The CDR1 regions lie between the helices and are able to contact both the MHC and the bound peptide (Garboczi et al. 1996; Garcia et al. 1998). This positioning ensures that the hypervariable CDR3 regions directly contact the peptide, while the less variable, germline-encoded, CDR1 and 2 regions interact with the more conserved MHC structure, see figure 1-1. The TCR α chain is thought to control TCR-MHC orientation and mediate most of the TCR's interaction with the MHC (review by Garboczi and Biddison 1999; review by Garcia et al. 1999).

Binding kinetics demonstrate a two-step binding process during TCR/MHC interaction and T-cell activation. The conserved MHC α-helical areas allow initial weak TCR engagement (CDR2, CDR1) regardless of the particular antigenic peptide identity, enabling the TCR CDR3 to scan the peptides of multiple MHCs (Manning et al. 1998; Wu et al. 2002). Positive CDR3/peptide interaction stabilises the binding and enables T-cell activation (Borg et al. 2005).

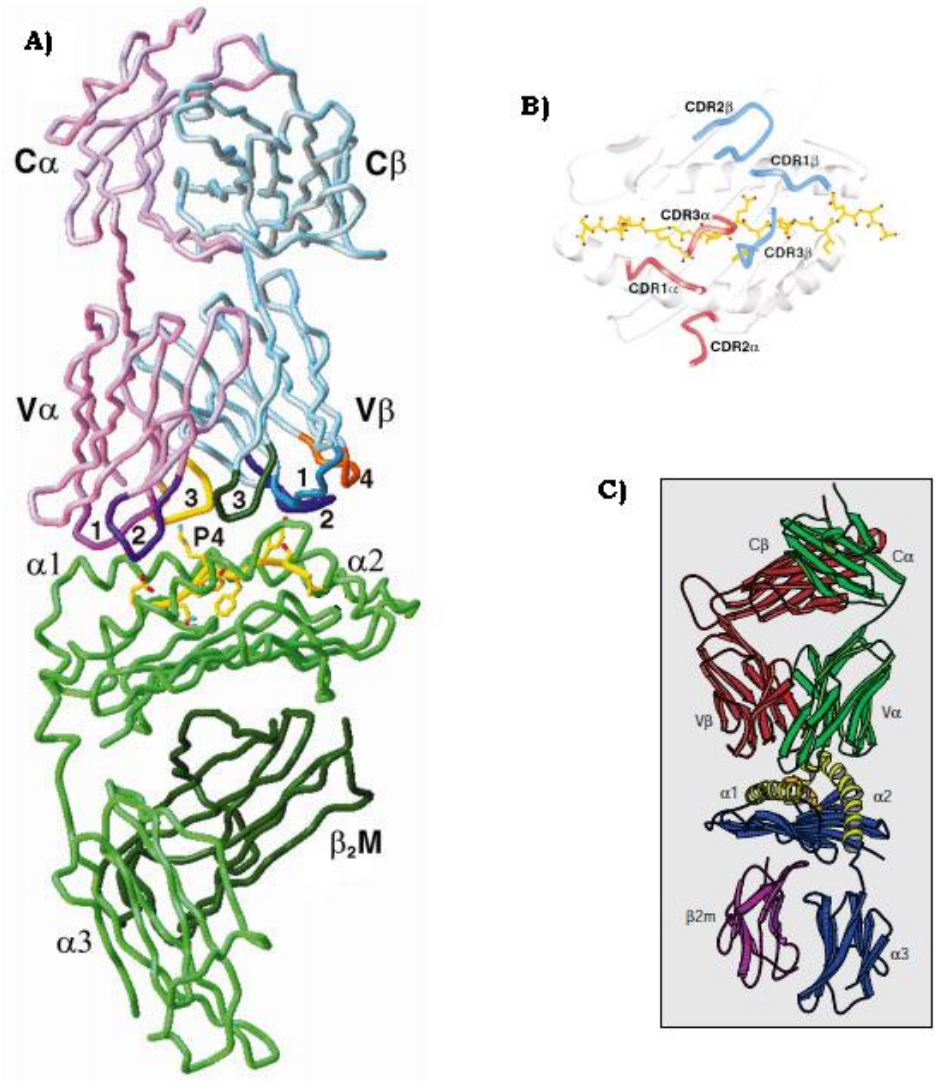


Figure 1-1: A) αβTCR/MHC-class-I structure and interaction. CDR regions are coloured: CDR1α magenta, CDR2α purple, CDR3α yellow, CDR1β cyan, CDR2β navy blue, CDR3β green, HV4 orange. The bound peptide (yellow) is central to the MHC α-helices (adapted from Garcia et al. 1998). B) CDR region placement over the peptide (yellow) and MHC α helices (grey). CDR2 bind to MHC only, CDR3 to peptide only, while CDR1 can interact with both (adapted from Hennecke and Wiley 2001). C) αβTCR/MHC-class-I interaction showing the predominance of the MHC α-helices (yellow) at the binding site. MHC-bound peptide is orange (adapted from Teng et al. 1998).

1.2.1.2 TCR Genomic Organisation

The TCRα locus is located on chromosome 14 in humans and mice (Boysen et al. 1997; Bosc and Lefranc 2003; Giudicelli et al. 2005), and chromosome 10 in cattle (Fries et al. 2001; Van Rhijn et al. 2007), and the genomic organisations of different species are relatively similar (Glusman et al. 2001). In each, the TCRδ genes are contained within the TCRα locus, both consisting of non-contiguous V, (D in the case of δ), J and C segments with interspersed disrupted (pseudogene) segments and locus-specific repeats. The α locus contains approximately 50 (human) and 104 (murine) Vα gene segments (including pseuedogenes) and 61 Jα gene segments (human). The δ locus is situated between the Vα and the Jα and Cα segments and in humans consists of 3 specific Vδ, 4 Dδ, and 4 Jδ gene segments and the four exons of the single Cδ gene. Gene rearrangement at this site is highly regulated, and TCRδ genes are rearranged early in thymic ontogeny. Later, α locus gene rearrangements result in the deletion of the δ locus. The C region genes for each chain consist of four separate exons which encode (i) the extracellular constant domain, (ii) a short 'hinge' region, (iii) the transmembrane and cytoplasmic regions, and (iv) 3' untranslated sequences (Satyanarayana et al. 1988; Takihara et al. 1988; Boysen et al. 1997; Scaviner and Lefranc 2000; Haynes and Wu 2004; Haynes and Wu 2007). The TCRα/δ locus is located adjacent to a number of olfactory receptor genes, two of which (OR10G2 and OR4E2) lie adjacent to the first Vα segments (between Vα7.1 and Vα7.2, and between Vα7.2 and Vα11.1, respectively) (Haynes and Wu 2007).

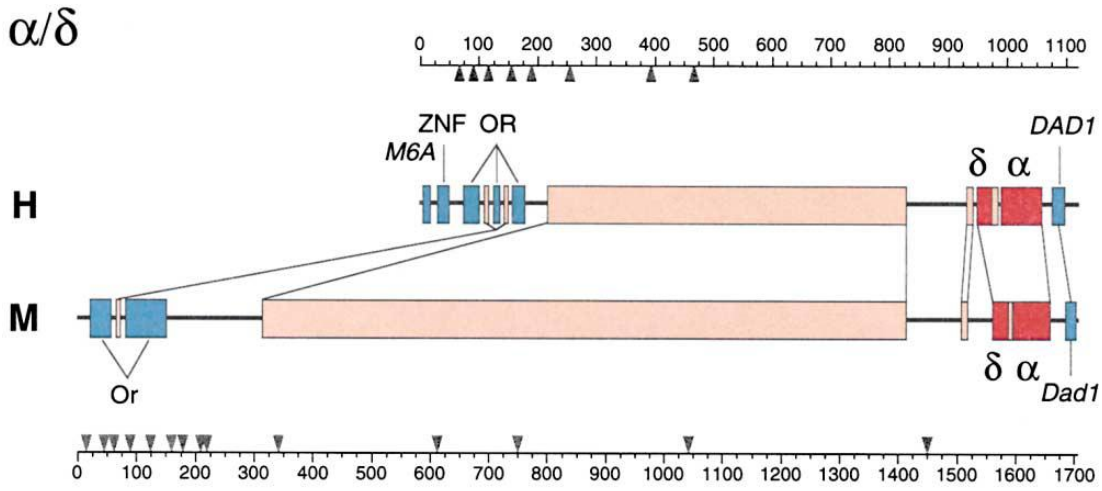


Figure 1-2: Overview of the TCRαδ locus (adapted from Glusman et al. 2001). The human (H) and murine (M) TCR loci are compared showing arrays of V segments in orange, other TCR elements in red, and non-TCR genes in blue. Gene symbols are indicated where available: M6A, a putative methyltransferase; ZNF, a zinc-finger protein; OR, olfactory receptor genes; DAD1, defender against cell death. Grey triangles indicate the locations of species-specific, processed pseudogenes.

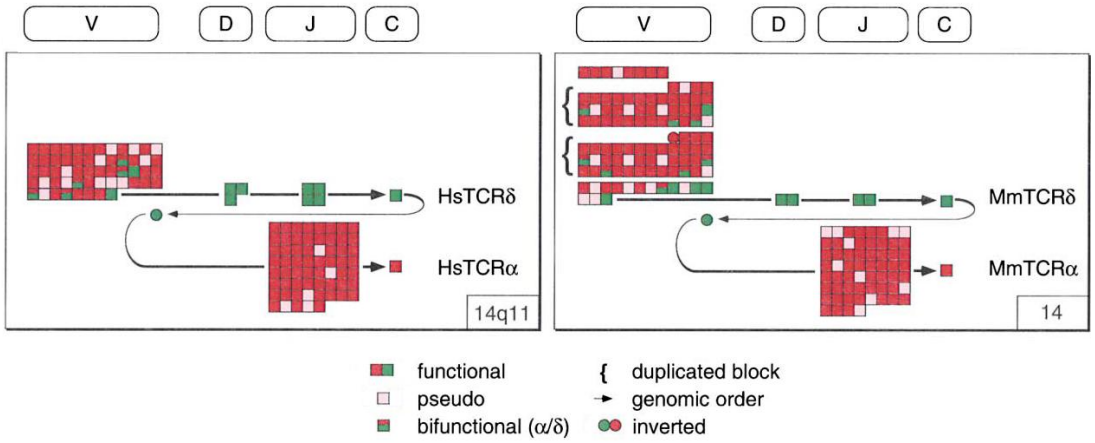


Figure 1-3: TCRαδ locus gene structure (adapted from Glusman et al. 2001). The genomic organisation of human (left) and murine (right) TCR genes, α (red) and δ (green), showing clusters of V, D, J and C gene segments aligned vertically.

1.2.1.3 TCR Nomenclature

TCR genes that are present as multiple gene segments in the genome (e.g. the V genes) were originally given numerical identification corresponding to their order of discovery. These sequences were classified into subfamilies, which in the case of the V genes included genes that shared over 70% nucleotide sequence identity. Comparisons between species have demonstrated that orthologues of many of the subfamilies are identifiable in the different species, indicating that the subfamilies arose prior to divergence of the species. Some subfamilies contain only a single member while others contain multiple members, the latter having arisen through later gene duplication events. Differences in the extent of gene duplication together with gene deletions in some species result in differences between species in the total complement of V genes. The World Health Organization – International Union of Immunological Societies (WHO-IUIS) published the first official nomenclature designation system for TCRs (Kazatchkine 1995). This was designed to unify the prefixes given to individual gene segments, but was complicated by the fact that the human and mouse TCR genes had been numbered independently, meaning that orthologous subfamilies rarely had corresponding subfamily numbers. Following sequencing of the human TCR genome loci, a second nomenclature system was introduced in 1999 by the creation of the IMmunoGeneTics (IMGT) database, http://imgt.cines.fr (Lefranc et al. 1998; Giudicelli and Lefranc 1999). This renamed the V subfamilies as subgroups and re-numbered them in the order in which they appear in the (human) genome locus, starting at the 5' end. This system was approved by the Human Genome Organization (HUGO) nomenclature committee and is now considered official (Glusman et al. 2001). In general, comparisons between species have shown a very similar organisation and order of V genes. However, because of differences between species in the number of V gene subgroups (due to gene deletions), this system still often results in orthologous subgroups being assigned different numbers. In species where the genome assemblies are incomplete (e.g. cattle) the V gene subgroups have generally been assigned numbers corresponding to their orthologues in humans (using either the IMGT or IUIS systems).

1.2.1.4 TCR genes in domestic animal species

TCRα genes equivalent to those in humans and mice have been described for a number of animal species, including horse (Schrenzel and Ferrick 1995), swine (Uenishi et al. 2003), cattle (Ishiguro et al. 1990) and sheep (Hein et al. 1991). The C genes show a relatively high degree of sequence conservation, and V gene subfamilies orthologous with those in humans are readily identifiable. Interestingly, although assembly of the bovine TCRαδ locus is incomplete, over 300 Vα and over 100 Vδ genes have already been identified. Analysis of the Vα gene sequences revealed 11 (of 41) human sub-families which are not represented in cattle and similarly 11 bovine sub-families which are not represented in the human genome (Reinink and Van Rhijn 2009). The overall 3-5 fold increase in Vα repertoire in comparison to mouse and human could potentially afford markedly increased diversity to the bovine αβ TCR repertoire.

1.2.2 TCR Diversity & Repertoire

Host defence in mammals relies on the ability of the adaptive immune system to recognise and respond to a diverse array of pathogens. Hence, the size and diversity of the available antigen-receptor repertoires, T-cell and B-cell, are crucial for the host immune defence.

1.2.2.1 T-cell development in the thymus

Lymphocytes differentiate from haematopoietic stem cells, early lymphoid progenitors within the bone marrow. Precursors of T-cells migrate to the thymus and undergo development including acquisition of TCR expression to generate the naïve T-cell population. Early thymocytes are negative for CD4 and CD8 and TCR

expression, but give rise to the major T-cell population within the thymus, which is double positive ($CD4^+/CD8^+$) and undergoes somatic rearrangement of TCR genes to achieve TCR expression. The TCR β locus is rearranged first, bringing together V, D, J and C gene segments to encode a TCRβ chain, which forms a complex with an invariant pre-TCRα chain and, together with CD3 polypeptides, enables low-level expression on the double positive thymocytes. Stimulation through this pre-TCR-CD3 complex causes these cells to cease β locus rearrangement and actively proliferate within the thymus, before initiating rearrangement of the α locus (Levelt et al. 1995; Xu et al. 1996). Rearrangement of α V, J and C gene segments results in α chains that can pair with the expressed β chains to form a TCR. This continual α locus rearrangement is aided by the locus structure, which allows multiple V/J recombination events on the same allele which result in the excision of the prior recombined DNA (Starr et al. 2003).

Cells expressing a TCR undergo rigorous selection for their ability to interact with MHC containing bound self peptides in the thymus. Cells unable to bind MHC-peptide die by neglect while those cells that bind to MHC-peptide with high avidity are also eliminated (Mathis and Benoist 2004). Less than 2% of thymocytes survive the selection process (Huseby et al. 2005). During this process, those cells that interact with class I MHC lose expression of CD4 and those interacting with class II MHC lose expression of CD8. These single-positive TCR^+ cells are exported from the thymus to form the naïve T-cell population (Goldrath and Bevan 1999; reviewed by Sebzda et al. 1999; Schlissel 2003; Starr et al. 2003; Turner et al. 2006).

1.2.2.2 Contribution of TCR gene re-arrangement to TCR diversity

The following elements are important in determining the extent of the TCR repertoire: **(i) Multiple germline gene segments** – the high number of different V (D) and J germline segments at each locus; **(ii) Combinatorial diversity of V (D) & J joining** – the massive number of possible different V (D) and J combinations for each chain; **(iii) Junctional diversity (& N-region diversification)** – random and

programmed nucleotide insertion, deletion and mutation at gene fusion sites on each chain; **(iv) Combinatorial association of random pairing of α & β chains** – the random pairing of both chains of the complete TCR (Allison and Lanier 1987; Davis and Bjorkman 1988; Glusman et al. 2001).

The generation of junctional diversity, which adds substantially to the TCR repertoire, is a consequence of the re-arrangement process. The rearranging TCR gene segments are flanked by recombination signal sequences (RSS), adjacent to the 3' end of V segments, the 5' end of J segments, and both ends of D elements in the β locus. These consist of highly conserved heptamer and nonamer motifs separated by a spacer of either 12 or 23 nucleotides (+/- 1). Gene segments of each particular type are all flanked by RSS with the same spacer length, and efficient recombination only occurs between RSS sequences of different spacer lengths, the '12/23 rule' (Akira et al. 1987). V(D)J recombination is initiated when the complex of lymphocyte-specific recombination-activating gene 1 (RAG1) and RAG2 binds to the RSS. These cleave the DNA at the gene-segment/RSS-heptamer junction, leaving 'coding ends' of covalently closed DNA hairpins, and blunt, phosphorylated 'signal ends' (McBlane et al. 1995; Livak and Schatz 1996). The promiscuously-expressed non-homologous end-joining (NHEJ) proteins, in combination with RAG1/RAG2, open the coding ends (by nuclease activity), then process and join these ends to form a 'coding joint'. This step is imprecise, and both nucleotide removal by deletion, and insertion via either palindromic duplications (P-segments) or non-templated nucleotide additions (N-regions, introduced by terminal deoxynucleotide transferase), can significantly alter the original germline sequence at this site (Cabaniols et al. 2001). The positioning of the CDR3 region across the V(D)J junction, at the site of this junctional diversity, ensures that it incorporates maximum sequence hypervariability. (reviewed by Schlissel 2003; Livak 2004; Schatz 2004)

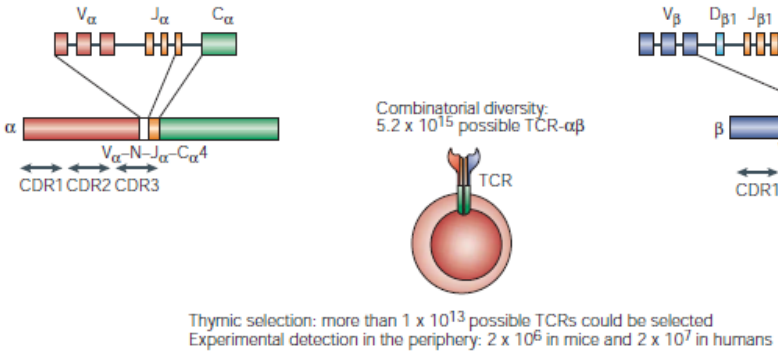


Figure 1-4: T-cell receptor αβ structural diversity (adapted from Nikolich-Zugich et al. 2004). Functional TCR chains are created by somatic V(D)J rearrangement from the TCR loci, of Variable (V) and Joining (J) genes for the TCRα chain, and V, Diversity (D), and J genes for the β chain. The random deletion, insertion or mutation of nucleotides at the site of gene joining (N-region) furthers increases the hypervariability of each chain sequence, particularly within the Complementarity Determining Region (CDR) 3. The massive diversity of the immature thymic T-cell population is vastly reduced by maturation, involving positive and negative selection within the thymus, to the peripheral naïve T-cell population.

The combination of different gene segments is not always random, especially within the TCRαδ locus, suggesting higher levels of control of V(D)J recombination. Rearrangement of specific V and J segments appears to be affected by several local factors, including control of access of the recombination 'machinery' to the TCR locus and the specific RSS sequences flanking gene segments (Krangel et al. 2000; Probst et al. 2004). Overall regulation mechanisms for TCR loci recombination are also apparent (Schlissel 2003): **(i) Lineage specificity** ensures recombination is only possible within the lymphoid lineage via the controlled expression of RAG proteins by transcriptional regulators. **(ii) Order within a lineage,** meaning that TCRβ rearrangement must precede TCRα rearrangement. **(iii) Allelic exclusion** ensures that a functional T-cell only generates one functional allele from each locus. This stage is less certain for the TCRα chain which undergoes processive rearrangement (see above). Although the presence of a productive TCRβ rearrangement prevents further β locus rearrangement (Uematsu et al. 1988), TCRα translation does not

strictly control α locus rearrangement, and evidence has been shown for some cell clones expressing two TCRα chains on the cell surface (Borgulya et al. 1992; Heath and Miller 1993; Padovan et al. 1993). Overall, α locus rearrangement terminates with the cessation of RAG expression associated with the positive selection of developing T-cells by thymic ligands.

The circulating T-cell repertoire has been estimated to contain 2×10^7 different TCRs in humans (Arstila et al. 1999), and 2×10^6 in mice (Casrouge et al. 2000). These estimates are much smaller than the theoretical mathematical estimate of 1×10^{15} TCRs obtained for human (Turner et al. 2006). This discrepancy is due both to limitations of different combinations to generate functional receptors and to the massive negative selection of rearranged T cells in the final stage of thymic development.

1.2.2.3 T-cell Repertoire Homeostasis

The pool of mature naïve T-cells and previously-activated memory T-cells circulate through the secondary lymphoid organs. The number of cells remains reasonably stable during health, although both populations undergo separate homeostasis and balanced loss and replacement. A number of different mechanisms are involved in T-cell homeostasis including controlling the rate of production, division of mature cells, trafficking and cell-death. These processes need to preserve not only the number of cells, but also both the immune system's 'memory' response to previously encountered pathogens and a wide diversity of antigen-receptors to new potential pathogens. Most naïve T-cells are maintained in the periphery without proliferating and survive for a number of weeks in the absence of antigen stimulation (Goldrath and Bevan 1999; Correia-Neves et al. 2001; Takada and Jameson 2009). In contrast, 'memory' T-cells are stable long-term, and display both qualitative and quantitative changes which ensure they can respond to repeat infections with enhanced kinetics (McHeyzer-Williams and Davis 1995; Savage et al. 1999; Blattman et al. 2000; Kedl et al. 2002).

Age-related changes in the T-cell repertoire cause a gradual decline in the number of naïve T-cells and increase in 'memory' phenotype T-cells, resulting in a narrowing of the TCR repertoire (Hosono et al. 1995; Sunderkotter et al. 1997). An attributing cause of this change has been hypothesised as the gradual involution of the thymus, which is particularly apparent following puberty, and decline in thymic export of naïve T-cells (Steffens et al. 2000). This loss of thymic activity does not affect the overall number of cells in the peripheral T-cell population, which undergoes thymus-independent homeostasis and is maintained by self-renewing memory cells and long-lived naïve T-cells (Lustig et al. 2009). However, some thymic function has been shown to continue, despite involution, throughout life, providing new naïve T-cells to the circulation (Cunningham et al. 2001; Dominguez-Gerpe and Rey-Mendez 2003). Inherent restrictions in possible TCRα locus V-J rearrangements of animals of different ages have also been hypothesised following a study which showed that V-J combinatorial diversity was restricted in both newborn and adult mice, yielding a trend from 5'-3' TCRα-J gene usage with age (Thompson et al. 1992).

1.2.3 The Major Histocompatibility Complex

The human MHC, a 3.6-Mb genetic region with over 100 genes, many of which have an immunological function, is one of four regions within the human genome that contain MHC-like genes. These are referred to as MHC-paralogous regions and they appear to have been created by two rounds of chromosomal duplication (polyploidisation) from a common ancestor (Abi-Rached et al. 2002). Current evidence indicates that approximately one third of MHC genes possess duplicated copies in, at least, one of the three paralogous regions. The MHC is located on chromosome 6, at 6p21.3-p22.2, and its three paralogous regions at 1q21-q25, 9q33-q34, and 19p13.1-p13.4, respectively. Of these, the 1q21-25 region is most phylogenetically related to the MHC region, and is the only paralog that contains true class I MHC-like genes, CD1 and MR1. The majority of genes identified in this chromosomal region also possess an immunological function, and it has been hypothesised that the 1q21-q25 region may represent another antigen-presentation

gene cluster (reviewed by Kasahara 1999; reviewed by Flajnik and Kasahara 2001; Shiina et al. 2001). Similar MHC paralogues have also been identified within the murine genome (Kasahara et al. 1996).

The human MHC is composed of three regions – class I, class II and class III. The class I and II regions contain a variety of highly-polymorphic genes within a well-conserved framework of gene clusters. The class I region contains 3 highly polymorphic classical class I genes, which encode the heavy chains of human leucocyte antigen (HLA)-A, -B and -C, as well as two clusters of non-classical class I genes, and various non-immunologically related genes. HLA-A has more than 80 alleles, HLA-B has over 180, while HLA-C has over 40. Both alleles of each locus are codominantly-expressed, resulting in the expression of up to six different MHC class I molecules (Gromme and Neefjes 2002). Non-classical class I genes generally show no or little polymorphism and in most cases their function is less well-defined than the classical class I genes. The class II region contains genes encoding the α and β chains of HLA-DM, -DP, -DQ, and -DR, as well as genes for TAP and components of proteasomal complex which are involved in the class I antigen processing pathway. The extended class II region also contains the gene encoding tapasin, a class I molecular chaperone (reviewed by Bjorkman and Parham 1990; reviewed by Beck and Trowsdale 2000; reviewed by Kelley et al. 2005). The Class III region, which in humans is a gene-dense area located between the class I and class II regions, contains a large number of structurally-unrelated genes, some of which encode proteins involved in the innate immune system, such as complement components and TNF-α (Yung Yu et al. 2000; Xie et al. 2003).

The extreme polymorphism of classical class I and class II MHC genes provides the capacity for selective pressures generated by repeated encounters with pathogens during evolution. Allelic variations in these genes primarily include nucleotide substitutions in those regions that encode the peptide binding groove of the MHC proteins (reviewed by Bjorkman and Parham 1990; reviewed by Beck and Trowsdale 2000).

The bovine MHC, or BoLA (Bovine Leucocyte Antigen), is sited on chromosome 23, and exhibits overall a similar organisation to that of the HLA. However, the bovine class II region has been split into two subregions by a chromosomal inversion, giving rise to one component containing class II genes of unknown function lying outside the MHC locus but on the same chromosome (Band et al. 1998). Whereas the human class I MHC has 3 expressed polymorphic classical class I genes, cattle have multiple (six or more) classical class I genes (Birch et al. 2006), but none of these is consistently expressed, and various combinations of between one and three different class I genes are expressed on different haplotypes (Ellis and Ballingall 1999; Ellis et al. 1999; reviewed by Ellis 2004). The combination of allelic diversity and diversity in gene number and combination affords greatly increased polymorphism to the cattle MHC repertoire; however, the significance of this increased diversity is as yet unclear.

1.2.3.1 Structure and function of MHC molecules

Class I and II MHC molecules function as antigen-presenting proteins, binding antigenic peptides that are presented to T-cells. The TCR recognises the entire MHC-peptide complex. Two subsets of T-cells distinguished by expression of the T-cell co-receptors CD4 and CD8 recognise antigens presented by class II and class I MHC proteins respectively.

Both MHC class I and class II molecules are heterodimers expressed on the cell surface. MHC class I molecules are composed of a heavy chain non-covalently associated with an invariant light chain, β-2-microglobulin (β2M), which is encoded by a gene lying outside the MHC. The heavy chain has an intracellular cytoplasmic tail and transmembrane region, while the extra-cellular portion is composed of three domains, α1, α2 and α3, each formed from discrete exons. The membrane-proximal α3 domain interacts with β2M, with which it is structurally homologous. The T-cell co-receptor, CD8, also interacts with the α3 domain during MHC – T-cell binding. The α1 and α2 domains contain two anti-parallel α-helices which form the sides of a

deep cleft that functions as the peptide-binding site. Typically MHC class I molecules present antigenic-peptides of 8-11 residues, which are bound in an extended conformation and include 2 or 3 residues that bind via side chains into pockets in the MHC groove.

The MHC class II proteins are heterodimers consisting of α and β chains which are non-covalently linked on the cell surface. Each chain has cytoplasmic and transmembrane domains and two extracellular domains. The membrane-distal domains of each chain (α1 and β1) combine to form a peptide binding groove similar to that of the class I. However, the ends of the groove are open, allowing the MHC class II molecule to present longer peptides (10-15 residues) in extended conformation (Maenaka and Jones 1999) (reviews by Bjorkman and Parham 1990; Bjorkman 1997; Rudolph et al. 2006; Takeshima and Aida 2006).

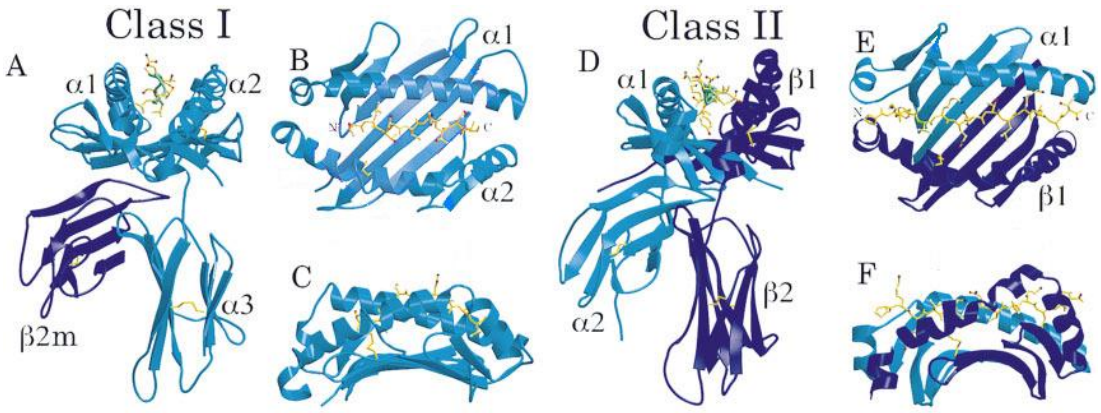


Figure 1-5: 3-dimensional structure of MHC class I and class II. (A)/(D) molecular composition of β-pleated sheets as well as α-helices forming the peptide-binding groove on the membrane-distal surface; (B)/(C)/(E)/(F): Close-up of the peptide-binding groove, formed by two antiparallel α-helices on a 'bed' of β-pleated sheets (of class I α1/α2 and class II α1/β1), and interaction with a peptide (adapted from Bjorkman 1997).

1.2.3.2 Antigen-Processing Pathways

MHC class I molecules are expressed on most nucleated cells and typically allow the immune system to monitor these cells for viral infection or neoplastic change by presentation of peptides derived from intracellular breakdown of cytosolic protein. To do this, the immune system utilises the cell's normal mechanisms of protein turnover and catabolism, including the proteasome complex, which degrades ubiquitinated proteins, and the action of aminopeptidases in the endoplasmic reticulum (ER). Peptides are translocated from the cytoplasm into the ER by a heterodimeric protein known as the transporter associated with antigen processing (TAP). The ER contains newly-synthesised MHC class I heavy chains which are assembled with β2M and initially associates with molecular chaperones, including tapasin and calreticulin. Along with TAP and other ER proteins, this forms the protein loading complex which aids optimal peptide loading of the MHC molecule. Successful peptide binding stimulates release of the MHC class I - peptide complex from the loading complex and transport via the Golgi apparatus to the cell surface (Rock 1996; Rock et al. 2002; reviewed by Yewdell et al. 2003; Rock et al. 2004; reviewed by Koch and Tampe 2006). This is a highly selective process which sees most of the oligopeptides generated from protein catabolism recycled and degraded within the cell rather than binding to MHC molecules (reviewed by Yewdell et al. 2003).

Expression of MHC class II molecules is mostly restricted to specialized antigen presenting cells such as macrophages, DC and B-cells. These molecules typically present exogenously-derived peptides which have undergone proteolysis within endosomal compartments. Exogenous antigens are endocytosed by these cells and digested by proteases, particularly cathepsins, into multiple small oligopeptides. MHC class II molecules synthesized within the ER of the cell are initially prevented from binding peptides by associating with a molecular chaperone, the class II invariant chain (Ii). This polypeptide occupies the peptide-binding groove and aids transport of the class II molecule, via the Golgi and *trans*-Golgi network, to late

endocytic vesicles. Here, endosomal cysteine proteases progressively degrade Ii until only the class-II-associated invariant-chain peptide (CLIP) remains in the peptide binding groove. This molecule is swapped for suitable antigenic peptide, under the control of another class II protein, DM, and the MHC-peptide complex is then transported to the cell-surface (Benaroch et al. 1995; Bryant and Ploegh 2004).

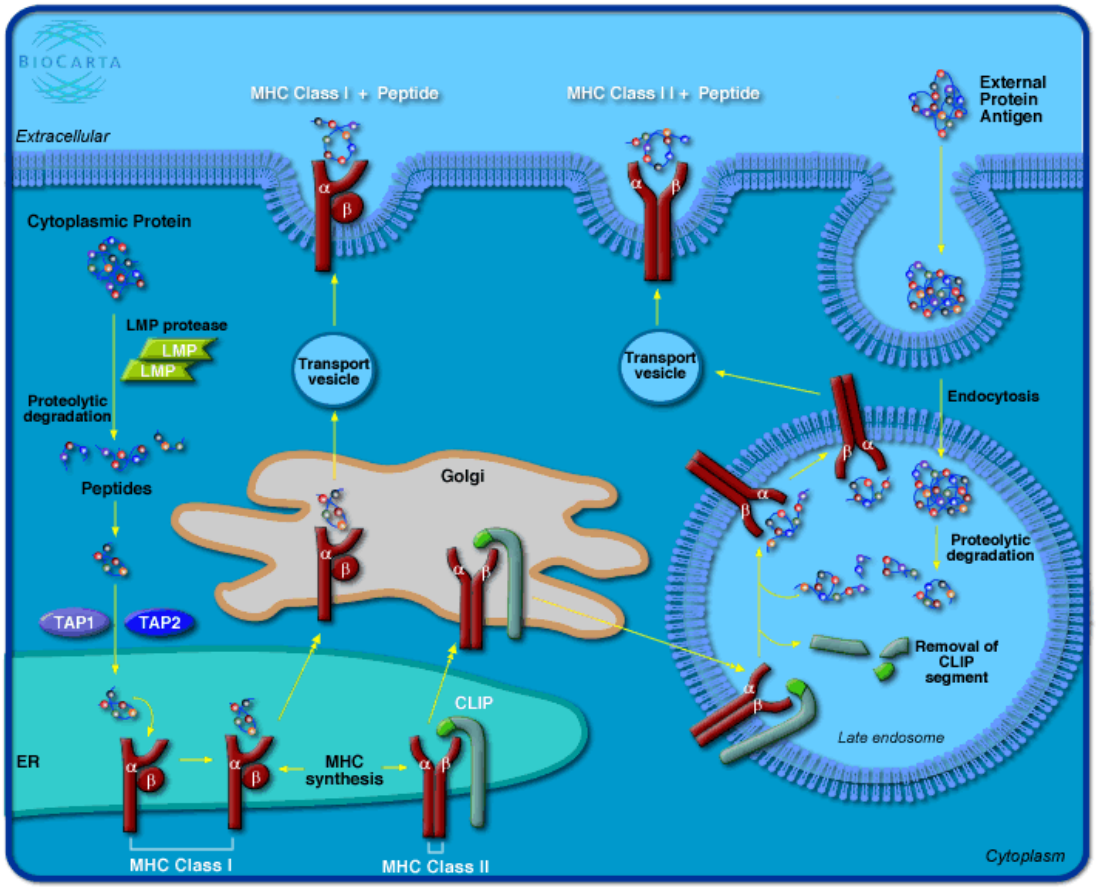


Figure 1-6: Comparison of the MHC class I pathway for presentation of endogenous protein, and the MHC class II pathway for presentation of exogenous antigen. (Image from http://kugi.kribb.re.kr/KUGI/Pathways/BioCarta/h_mhcPathway/, Korean Unigene database, December 2009).

1.2.3.3 Non-classical class I MHC molecules

In contrast to classical class I genes, non-classical class I MHC genes (or class Ib genes) exhibit limited or no polymorphism and tend to be expressed on a more limited number of cell types (Fischer Lindahl 1993). The MHC class Ib molecules also include a number of non-MHC-encoded class I-like proteins, such as the CD1 family, the neonatal Fc receptor FcRn, and MR1 (reviewed by Braud et al. 1999).

The function of several of the human non-classical class I molecules has been identified. HLA-E, which is ubiquitously expressed, has been shown to regulate NK cell function by interacting with the cell-surface CD94/NKG2 receptors. It also presents hydrophobic peptides to a small subset of $CD8^+$ T cells known as NK-CTL, which exhibit non-MHC-restricted killing. HLA-G is expressed at the maternal/foetal barrier and is hypothesised to regulate the immune response against the foetus. HFE (formerly HLA-H) has a narrowed peptide-binding groove and performs a non-immunological function in iron metabolism. Other non-classical molecules have a less similar structure to classical class I MHC, including lack of requirement for β2M. Zn-α2-glycoprotein (ZAG) has a large binding groove and is potentially involved in lipid catabolism. The MICA/MICB genes are encoded in the MHC class I region but have no function in antigen presentation and have been identified as general cellular-stress sensors – these encode proteins which act as ligands for the NKG2D type II receptor and stimulate the cytolytic response of NK cells, $CD8^+$ αβ T-cells and γδ T-cells expressing this receptor (reviewed by Braud et al. 1999; Maenaka and Jones 1999; Steinle et al. 2001; Stephens 2001; reviewed by Rudolph et al. 2006; Pietra et al. 2009).

The class I-like genes in humans, encoded outside the MHC, include the CD1 family and MR1, which map to the MHC-paralogous region on chromosome 1. Both display remarkably high conservation of the α1 and α2 domains across mammalian species. Human CD1 molecules are characterised by their ability to present non-protein antigen to T-cells, due to their deep, narrow, hydrophobic binding groove (Kawano et al. 1997). CD1a, CD1b and CD1c present bacterial cell wall lipids primarily to $CD8^+$ and $CD4^-$ $CD8^-$ αβ T-cells, while CD1d, which is expressed on all

bone-marrow-derived cells, also binds lipids. CD1d (Godfrey et al. 2000) and MR1 (Treiner et al. 2003) molecules have been identified as the restriction elements for two distinct subsets of 'invariant' T-cells, invariant Natural Killer T (iNKT) cells and Mucosal Associated Invariant T (MAIT) cells, respectively (see section 1.3).

Functional counterparts of HLA-E, FcRn, HFE, the CD1 family, including CD1d, and MR1 have also been identified in mice (Braud et al. 1999) and in the case of CD1d and MR1 a high degree of conservation in sequence is observed between the species, suggesting that they have an evolutionarily-conserved function (Blumberg 1998; reviewed by Braud et al. 1999; Treiner et al. 2003; reviewed by Rudolph et al. 2006).

In cattle, in addition to the classical class I genes, the MHC contains four MIC genes, and four other non-classical class I genes of unknown function, three near the MIC genes and one near the classical class I region, as well as a number of non-classical psuedogenes (Birch et al. 2008; Birch et al. 2008). Outwith the MHC, CD1 genes have also been identified, but to date a functional CD1d gene has not been identified (Van Rhijn et al. 2006).

1.3 T-cell subsets expressing invariant TCRα chains

T-cell populations characterised as bearing invariant TCRα chains are iNKT cells and MAIT cells. The TCRα chains of these T cells were first identified as occurring at a higher than expected frequency among sequences of human TCRα cDNAs. This led to the identification two distinct subsets of T-cells known as invariant NK T-cells (iNKT) and mucosal associated invariant T-cells (MAIT) (Porcelli et al. 1993). T-cell subsets expressing invariant TCRα chains encoded by the orthologous V and J gene segments have also been identified in mice (Tilloy et al. 1999).

1.3.1 Invariant NKT-cells and CD1d

The human iNKT-cell TCRα chain utilises the Vα24 and Jα18 gene segments and has a highly conserved CDR sequence; this chain is frequently paired with a TCRβ chain utilising the Vβ11 gene segment (Dellabona et al. 1994). The corresponding invariant murine TCRα chain utilises the orthologous V and J gene segments (Vα14 and Jα18) and preferentially pairs with TCRβ chains utilising the Vβ8, Vβ7, or Vβ2 gene segments (most frequently Vβ8.2) (Lantz and Bendelac 1994). The cells in mice were named NKT cells because they were originally identified as expressing the NK cell marker CD161/NK1.1 (Makino et al. 1995).. These invariant T-cells were later shown to be restricted by the non-classical MHC Ib molecule CD1d (Godfrey et al. 2000), in a β2M-dependent, TAP-independent manner. Accordingly, the term 'NKT-cells' has now come to encompass the CD1d-restricted T-cell population. The $CD161^{+}$/$NK1.1^{+}$ T cells were subsequently shown to contain subsets not restricted by CD1d and with diverse TCRα repertoires, which were re-named 'NKT-like cells' (Gumperz and Brenner 2001; Godfrey et al. 2004). Human and murine NKT cells have recently been re-classified as either 'classical'/'type I' NKT-cells, (or 'invariant NKT-cells', iNKT), with invariant TCRα and restricted TCRβ repertoire, or 'non-classical'/'type II' NKT-cells, which are CD1d-restricted but do not use the invariant TCRα chain (Seino and Taniguchi 2005; Roy et al. 2008). In both species, iNKT cells have been further defined into $CD4^{+}$ and $CD4^{-}$

(either $CD4^-CD8^+$ or $CD4^-CD8^-$) populations, which appear to have distinct functions (Seino and Taniguchi 2005).

Both human and murine iNKT-cells recognise glycolipid antigens, as well as bacterial lipids, in association with CD1d molecules (Kawano et al. 1997; Zhou et al. 2004; Fox et al. 2009). Activation of iNKT-cells results in rapid cytokine release – in humans; $CD4^+$ iNKT cells produce both Th1 and Th2 cytokines, such as IL-4, IL-10, and IFN-γ, while $CD4^-$ iNKT cells selectively produce the Th1 cytokines IFN-γ, TNF-α, and also upregulate perforin (Gumperz et al. 2002; Lin et al. 2006). iNKT cells have been shown to mediate diverse protective and regulatory immune functions, including tumour rejection, protection against infectious microbes, and suppression of autoimmunity, by altering the Th1/Th2 balance (reviewed by Linsen et al. 2005; Seino and Taniguchi 2005; Kim et al. 2008). Activated iNKT cells also induce DC and B-cell maturation, and NK cell activation, effectively bridging the adaptive and innate immune responses (reviewed by Cerundolo et al. 2009).

iNKT cells are more abundant in mice than in humans (Treiner and Lantz 2006), accounting for <1% of mouse peripheral T cells (and up to 50% of liver and bone marrow T-cells) (Bendelac 1995; Ishihara et al. 1999) compared to 0.2% (0.01% - 0.92%) of human peripheral blood T-cells (Linsen et al. 2005; Montoya et al. 2007). Full-length iNKT cell TCRα transcripts displaying a high level of similarity to the human and mouse sequences have been identified in other species, including horse, pig, rabbit, and African elephant. However, functional CD1d appears absent from some ruminant species, including *Bos taurus* cattle. Instead, two CD1d psuedogenes are formed with mutations in the start codon and first intron (Van Rhijn et al. 2006; Looringh van Beeck et al. 2009). The lack of a functional CD1d gene may suggest the absence of iNKT cells in these species, or that another gene product performs the equivalent function.

Table 1-1: Comparison of the main features of CD1d- and MR1-restricted T-cells (adapted from Treiner and Lantz 2006)

Main features of CD1d- and MR1-restricted T cells		
	CD1d-restricted (NKT) cells	MR1-restricted (MAIT) cells
TCRα	Mainly invariant mVα14/hVα24-Jα18	Semi-invariant mVα19/hVα7.2-Jα33
TCRβ	mVβ2,7,8/hVβ11	mVβ6,8/hVβ2,13
Phenotype	DN or CD4 (or human CD8αα)	Mainly DN (or human CD8αα)
Frequency	High in mice, low in humans	Low in mice, high in humans
Anatomic location	Liver, thymus, spleen	Gut LP
Selecting/recruiting cells	Cortical thymocytes	Haematopoietic cells
Antigen-presenting cells in periphery	DCs, B cells, activated T cells, macrophages, hepatocytes	B cells
Lymphokines	IL-4, IL-13, IFNγ, IL-10, TGFβ	IL-4, IL-5, IL-10, IFNγ
Ligands	Endogenous and exogenous lipids	Endogenous and/or exogenous hydrophilic molecules?
Pathologies involved in	Infections, autoimmunity, cancer, inflammation	Central nervous system inflammation
Functions	Immune regulation?	Mucosal immune regulation?

1.3.2 MAIT Cells and MR1

MAIT cells are characterised by expression of an invariant TCRα chain, and their restriction by the MHC class I-like molecule, MR1. They have been described in mice, humans and cattle and the TCRα sequence is highly conserved between the species. The MAIT rearranged TCRα gene utilises the orthologous V and J gene segments in each species - murine Vα19-Jα33, human Vα7.2-Jα33 and bovine Vα19-Jα33. They also have a highly conserved CDR3 region of constant length (Tilloy et al. 1999; Treiner et al. 2003). Interestingly, the MAIT TCRα V gene is located at the immediate 5' end of the TCRVα locus and is interspersed between olfactory receptor genes; its presence as the first (mice, cattle) or second (human) Vα gene in the locus appears to be consistent between species (Reinink and Van Rhijn 2009). The MAIT TCRα chain is preferentially associated with TCRβ chains containing a limited repertoire of V genes – Vβ6 and Vβ8 in mice, and Vβ2 and Vβ13 in humans. The repeated use of the same TCRα and β junctional re-arrangements in different cells from the same subject, and the fact that these were absent in other subjects, demonstrated the oligoclonal nature of this repertoire, relating to the peripheral clonal expansion of circulating MAIT cells. (Porcelli et al. 1993; Tilloy et al. 1999; Treiner et al. 2005).

MAIT TCRα transcripts were originally demonstrated to be abundantly expressed in the $CD4^{-}CD8^{-}$ DN T-cell population of humans, mice and cattle. MAIT TCRα transcripts were measured by quantitative PCR, and quantified in relation to total TCRα measured using C-region specific primers. MAIT TCRα transcripts were also detected at significant levels in the human $CD8\alpha\alpha^{+}$ and the murine $CD4^{+}$ T-cell populations. MAIT TCRα transcripts were not detected at significant levels in bovine $CD4^{+}$ and $CD8^{+}$ T-cells. However, sequencing of multiple cDNAs revealed that the invariant Vα19Jα33 chain was abundant in bovine DN T-cells (Tilloy et al. 1999). A more recent publication studied MAIT cells by using monoclonal antibody against the Vα7.2 chain; since this antibody was not exclusively specific for MAIT TCRα (but instead all T-cells bearing Vα7.2), all results were confirmed by sequencing of the full TCRα chain. This study demonstrated abundance of MAIT cells within the human $CD8\alpha\beta^{+}$ population. Within $CD3^{+}/\gamma\delta^{-}$ lymphocytes, the study found that MAIT cells represented 40% of DN/$CD8\alpha\alpha^{+}$ T-cells, 8% of $CD8\alpha\beta^{+}$ T-cells, but <1% of $CD4^{+}$ T-cells (Martin et al. 2009). Correspondingly, TCRα CDR3 spectratype analysis of $CD4^{+}$, $CD8^{+}$, and DN subsets of human T-cells, as well as overall peripheral blood mononuclear cells (PBMC), revealed that the invariant Vα7.2 CDR3 sequence was over-represented in all αβ T-cell subsets; this was highest in the DN and $CD8^{+}$ subsets, at 4.3 or 3.5 times expected frequency, respectively (Han et al. 1999).

MAIT cells are 5-10 times more abundant in humans than mice (Treiner et al. 2005). Kinetic ELISA PCR study demonstrated them as accounting for 0.1-0.2% of all human peripheral T-cells (Tilloy et al. 1999); however, a definitive murine figure for comparison has not been published. Further kinetic ELISA PCR studies showed that MAIT cells represent 15% of human circulating DN T-cells, but only 2% of murine lymph node DN T-cells (Tilloy et al. 1999), although these figures cannot allow direct comparison between species. More recently, the use of monoclonal antibody against human Vα7.2 suggested that MAIT cells represented as much as 1-4% of human peripheral T-cells (Martin et al. 2009).

Analyses by quantitative PCR originally demonstrated high levels of MAIT TCR transcripts in PBMC and lymph nodes. Subsequent analyses of T cell subsets purified by flow cytometry demonstrated that MAIT TCRα transcripts were highly abundant in intestinal mucosa and were enriched in cells expressing the gut-specific α4β7 integrin. They were particularly prominent in the intestinal lamina propria, and to a lesser extent within the Peyer's patches and mesenteric lymph nodes; virtually no MAIT cells were detected within intestinal epithelium. Expression within the lamina propria was roughly ten-times that within the mesenteric lymph node. MAIT TCRα transcripts have also been detected in human and murine lung tissue (Tilloy et al. 1999; Treiner et al. 2003; Treiner et al. 2005).

1.3.2.1 MAIT Cell Development and Phenotype

Initial studies had shown that MAIT cell development was dependent on a thymus, as MAIT cells were absent in athymic *nude* mice (Tilloy et al. 1999), and that their selection was dependent on interaction with ligands expressed on haematopoietic-derived cells (rather than thymic epithelium) (Treiner et al. 2005). MAIT cells were also shown to be notably fewer in MR1$^{-/-}$ knock-out mice, suggesting that they were selected and/or restricted by MR1 (Treiner et al. 2003). However, although MAIT cells were not detected within cord lymphocytes, thymus or spleen of neonatal animals, large numbers of MAIT cells were detected in peripheral lymphoid organs of older animals, and these demonstrated an oligoclonal Vβ repertoire suggesting peripheral population expansion (Tilloy et al. 1999). Additionally, MAIT cells were also absent from both B-cell deficient humans and mice, as well as from gnotobiotic mice, suggesting that their selection and/or expansion was dependent on B-cells and commensal flora (Treiner et al. 2003).

These findings were clarified by a recent study using antibody against human Vα7.2, which showed that MAIT cell development is a stepwise process of initial thymic selection followed by peripheral expansion (Martin et al. 2009). Intra-thymic MAIT cells exist only in low numbers and have a naïve phenotype. The positive effect of

MR1 on foetal thymic organ culture and double iTCRα and TCRβ transgenic RAG knockout mice demonstrated that thymic selection of MAIT cells was dependent on both MR1, and a non-B-cell, non-T-cell, haemopoietic-derived cell-type. Since MR1 was shown to be absent from thymic epithelial cells, the outstanding likelihood is that this represents MR1 expression on thymic dendritic cells or macrophages. This thymic selection was neither dependent on B cells nor on commensal bacterial flora, whereas the subsequent large-scale peripheral expansion of MAIT cells was dependent on both B-cells and commensal flora. This process was confirmed by adding back B cells (expressing MR1) to mice engineered to only express MAIT TCRα at the surface of all T-cells; this resulted in a 10-fold increase in MAIT cell population. Interestingly, B cells lacking the MR1 molecule were also able to stimulate MAIT cell expansion (to a lesser degree), suggesting that MR1/TCR interaction in the periphery may not be necessary for MAIT development (Treiner et al. 2005; Martin et al. 2009).

As with NKT cells, the majority of MAIT cells express NK cell surface markers, particularly CD161 (NKR-P1A), which was present on the large majority of human blood MAIT cells detected using a monoclonal antibody specific for TCR Vα7.2. In contrast, NK1.1 was detected on less than 30-50% of murine MAIT cells (Martin et al. 2009). A number of studies have shown that circulating human MAIT cells express a predominantly memory T cell surface phenotype: ($CD45RO^{+}$, $CD45RA^{-}$, $CD56^{-}$, $CD57^{-}$, $CD28^{+}$, $CD27^{+}$ (Tilloy et al. 1999); $CD45RA^{-}$, $CD27^{+}$, $CD44^{hi}$, $CD28^{+}$, $CD161/NKR\text{-}P1A^{+}$, $\alpha4\beta7^{+}$, $CD56^{-}$, $CD57^{-}$ (Treiner et al. 2005); $CD45RO^{+}$, $CD45RA^{-}$, $CD27^{+}$, $CD62L^{lo}$, $CCR7^{-}$, $CXCR6^{+}$ (Martin et al. 2009)). Interestingly, human cord blood MAIT cells display a naïve T-cell phenotype, $CD45RA^{hi}$, $CD27^{hi}$, $CD45RO^{lo}$, although the majority are $CD161^{+}$ at this stage. This suggests that these cells acquire CD161 expression within the thymus and retain a naïve phenotype until they undergo peripheral expansion (Martin et al. 2009).

1.3.2.2 MR1 as a restriction element for MAIT cells

The non-classical class I gene MR1 is a single copy gene located in the MHC-paralogous region on human chromosome 1q25.3 and in mice at chromosome 1H1, near the CD1 family (Hashimoto et al. 1995; Yamaguchi et al. 1997). Orthologous MR1 genes have been detected in a number of other species, including rat and non-human primates (Walter and Gunther 1998; Parra-Cuadrado et al. 2001). The sequences of MR1 in different species show a high level of conservation both at the nucleotide and amino acid levels (human and mouse MR1 show around 90% amino-acid identity of the α1 and α2 domains) (see table 1-2). This inter-species similarity is the highest observed for any non-classical MHC class I molecule, and suggests an evolutionarily-conserved function.

Table 1-2: Percentage similarity of amino-acid and nucleotide (in parentheses) sequences of various MHC class I-related molecules between human and mouse (adapted from Yamaguchi et al. 1997).

Human / mouse	L	α1	α2	α3	TM&CY
MR1(h) / MR1(m)	40.9 (59.1)	89.7 (85.8)	89.1 (87.0)	72.8 (81.5)	40.7 (59.0)
HLA-A2 / H2-k^d	62.5 (62.5)	68.9 (79.6)	76.1 (81.9)	76.1 (83.0)	37.7 (60.2)
HLA-H / MR2	34.5 (47.1)	78.4 (79.2)	70.0 (73.7)	69.6 (74.6)	51.0 (71.2)
FcRn(h) / FcRn(m)	50.0 (58.3)	70.9 (70.5)	65.2 (71.7)	73.3 (72.2)	57.3 (64.0)
CD1d(h) / CD1d1(m)	60.0 (70.0)	59.3 (74.4)	62.4 (77.1)	76.3 (82.4)	25.0 (53.7)
Znα2gp(h) / Znα2gp(m)	77.3 (75.6)	58.6 (73.6)	73.9 (77.9)	46.2 (69.2)	

L: Leader peptide, TM: Trans-membrane domain, CY: Cytoplasmic domain.

Despite lying outside the MHC, the MR1 gene bears a high degree of sequence-similarity to classical class I MHC genes – the highest of any non-classical MHC class Ib molecule. The MR1 heavy chain associates with β2M and the α1 and α2 domains are predicted to form a groove similar to the peptide-binding groove of

classical class I proteins (Hashimoto et al. 1995; Yamaguchi et al. 1997; Yamaguchi and Hashimoto 2002; Miley et al. 2003). MAIT cells were shown to be absent in β2M-deficient mice (Tilloy et al. 1999).

Studies using PCR and Northern blotting have demonstrated ubiquitous transcription of MR1 amongst different murine organs and human cell-lines of various different lineages (Hashimoto et al. 1995; Riegert et al. 1998). This revealed a number of differently-sized transcripts, which were later shown by sequencing to represent alternatively spliced isoforms in both humans and mice. Two of the three identified human alternatively-spliced transcripts exhibited omission of the entire α3 region while retaining the open reading frame; the third transcript exhibited in-frame truncation of the α3 region. The five murine isoforms contained in-frame truncation or omission of the α1 region. The α2 domain was conserved in all splice variants, except one murine isoform which contained an intron insert (Riegert et al. 1998).

Cell surface expression of MR1 has yet to be demonstrated consistently in *ex vivo* tissues or in cell lines, and has been difficult to achieve in cells transfected with MR1 *in vitro*. In transfected P388 and L cell lines, MR1 protein was produced, but was retained intracellularly. However, human Hela cells and B6/WT-3 cells transduced with MR1 showed low levels of expression of the molecule at the cell surface. By using antibodies shown to be specific for open and folded configurations of MR1, both forms were detected on the cell surface and evidence was obtained that only MR1 in the folded conformation was able to activate MAIT cell hybridomas. This led to the hypothesis that MR1 is dependent on endogenous ligand for folding and surface expression (Huang et al. 2005; Treiner et al. 2005; Hansen et al. 2007).

Since MAIT cells fail to undergo peripheral expansion in B-cell-deficient mice (Treiner et al. 2003), it has been suggested that MR1 is presented to MAIT cells on the surface of B-cells, although this has not been proven. The failure to expand MAIT cells in B-cell deficient mice has been further linked to the B2 B-cell subset which include the cells that produce IgA in the gut (Treiner et al. 2005; Huang et al. 2008). Other recent work has supported the link with B cells, demonstrating low

MR1 expression on some human lymphoid cell lines, and particularly a subpopulation of $CD38^+$ $CD138^+$ and IgA^+ plasma cells or plasmoblasts derived from human gut mucosa (Gozalbo-Lopez et al. 2009).

1.3.2.3 MR1 Antigen-Presentation to MAIT Cells

1.3.2.3.1 Putative Ligands

Although there is evidence that MR1 activates MAIT-cells in a ligand-dependent manner (Hansen et al. 2007), the nature of this ligand has yet to be determined. Previous studies (section 1.3.2.2) have suggested that cell-surface expression of MR1 may be controlled by a ubiquitously-expressed endogenous ligand. Further, it has been hypothesised that the commensal flora in the gut may either provide exogenous or induce endogenous ligands for MR1, or induce translocation of MR1 to the cell surface via a cellular 'stress' signal (Treiner et al. 2003; Illes et al. 2004; Huang et al. 2005; Treiner et al. 2005; Croxford et al. 2006; Hansen et al. 2007; Huang et al. 2008).

The α1 and α2 domains of MHC class I and class I-like molecules define the binding groove of these molecules, and hence their ability to bind ligands. This is specifically reflected in the conserved residues involved in anchoring peptide termini, particularly the A and F pockets of classical class I molecules (formed by five and four residues respectively) which allow binding to peptide-chain amino and carboxyl termini (Bouvier and Wiley 1994). This region varies dramatically in different class I-like molecules, enabling them to take on unique functions. MR1, however, has one notable change and five conservative substitutions within these residues, resulting in loss of typical A- and F-pocket chemistry. Hence it is unlikely that MR1 binds a protein ligand, or if so it must use a different binding paradigm to classical class I molecules (Miley et al. 2003; Hansen et al. 2007).

Although molecular modelling suggests the ligand-binding groove of MR1 is hydrophilic (Treiner and Lantz 2006), the partially-hydrophobic synthetic glycolipid α-mannosyl ceramide (α-ManCer) has been shown to activate MAIT cells (Okamoto et al. 2005). This compares with the activation of iNKT cells by synthetic α-GalCer (Kawano et al. 1997; Pellicci et al. 2009). α-ManCer was a potent stimulator for MAIT cell activity, and minor alterations in the structure of its glycolipid chain affected the Th1-Th2 polarisation of the cytokine response (Shimamura et al. 2007). These findings suggest the likelihood that, as with CD1d and iNKT cells, MR1 presents lipid antigens to MAIT cells (Schumann and De Libero 2007).

1.3.2.3.2 MR1 Antigen-Presentation Pathways

The pathways through which MR1 presents ligand/antigen to MAIT cells are unclear. MR1 was shown to associate with the peptide-loading complex in the endoplasmic reticulum, but MAIT cell development appears to be independent of TAP and Ii, as MAIT cells are present in animals deficient in these molecules (Tilloy et al. 1999; Miley et al. 2003; Treiner et al. 2003). Unusually, despite its similar molecular structure to MHC class I molecules, transport of MR1 has been shown to follow an endocytic pathway. In this pathway, MR1 is transported from the Golgi to a late endosomal compartment before cell surface expression. Experimental inhibition of vacuolar acidification of the endosomal compartments and of endogenous Ii expression both dramatically inhibited MAIT cell activation by MR1; however, this did not reduce MR1 surface expression. It is hypothesised that MR1 may leave the ER when associated with a chaperone (possibly Ii), which traffics it to the late endosomal compartment to obtain a ligand for surface expression. This may include the 'recycling' of MR1 from cell-surface to endosomes as seen with CD1 molecules (Huang et al. 2008).

1.3.2.4 Activation and Function

MAIT cells produce a rapid and diverse cytokine response upon activation. This is predominantly a Th2 response including IL-4, IL-5, and IL-10, but can also include IFNγ and TGFβ. Activation of MAIT cells then additionally triggers the release of pro-inflammatory cytokines from bystander cells. The role of MAIT cells in regulation of (autoimmune) inflammation is suggested by their ability to polarise their cytokine response, particularly towards Th2; similar polarisation has been demonstrated in iNKT cells (Kawachi et al. 2006). When MAIT cells are over-represented, or when the intensity and duration of MAIT cell TCR-engagement is increased, the ratio of Th2 cytokines such as IL-4, IL-5 and IL-10 is markedly increased in comparison to the production of pro-inflammatory cytokines IFNγ and IL-17 (Croxford et al. 2006; Shimamura et al. 2008). This mimics the ability of different derivatives of the exogenous ligand α-ManCer to polarise the Th1/Th2 cytokine response of MAIT cells (Schumann and De Libero 2007). The heterogenous cytokine production in mice has been attributed to two potentially different groups of MAIT cells – one which produces IFNγ, IL-4 and IL-5 after TCR ligation, and another, predominantly $NK1.1^-$, which produces IL-10 and the costimulatory molecule ICOS (Kawachi et al. 2006; Hansen et al. 2007; Peterfalvi et al. 2008).

MAIT cells have been shown to be involved in the pathology of central nervous system inflammation, and accumulate in human multiple sclerosis (MS) lesions, as well as in the cerebrospinal fluid of MS patients. Development of experimental autoimmune encephalitis, a mouse model of MS, was shown to be inhibited in transgenic mice over-expressing T-cells bearing the MAIT TCRα, and exacerbated in MR1-deficient mice. MAIT cells have also been discovered in peripheral nerve biopsies from humans with chronic inflammatory demyelinating polyneuropathy. The authors of these studies proposed that MAIT cells were acting to reduce inflammation in these lesions through the release of Th2 cytokines, particularly IL-4 and IL-10 (Illes et al. 2000; Illes et al. 2004; Croxford et al. 2006). MAIT TCRα expression was also detected within human kidney and brain tumours at much higher levels than in the peripheral blood (Peterfalvi et al. 2008).

In addition to their proposed general immunoregulatory function, it has been suggested that MAIT cells may play a positive role in the homeostasis of intestinal mucosal immunity. Since mucosal tissues provide the largest area of contact between the body and the outside world, MAIT cells may play a role in the determination of commensal versus pathogenic bacteria on the mucosal surface, or perhaps in the production, or feedback control, of the large levels of protective IgA on these surfaces (Huang et al. 2005; Treiner et al. 2005; Treiner and Lantz 2006; Huang et al. 2008).

1.4 Aims of the Project

The purpose of this project was to study expression of the orthologous MAIT TCRα chain in cattle and sheep, to provide a marker that would allow further characterisation of MAIT cells in these species with respect to their tissue distribution, phenotype and potential function. This study also aimed to identify the putative restriction element for MAIT cells, MR1, and examine its expression both *in vivo* and *in vitro*. The specific objectives of the project were:

- To identify the orthologous MAIT TCRα in cattle and sheep and analyse cross-species conservation of the gene sequence. Furthermore, to determine the frequency of use of each germline element of the MAIT TCRα within the bovine TCRα repertoire.

- To develop a robust qPCR assay to provide a quantitative estimate of MAIT cell number in T-cell populations in order to examine the tissue distribution and phenotype of these cells.

- To identify the putative restriction element, MR1, in cattle and sheep and to study the cross-species conservation of this gene.

- To investigate the expression of bovine MR1, both with regards to tissue distribution of transcript, and surface expression of MR1 protein.

2 Materials and Methods

2.1 Experimental animals

Bovine peripheral blood was collected from ten adult Friesian-Holstein *Bos taurus* cattle. Ovine peripheral blood was collected from adult Blackface-cross ewes. Cattle blood and tissue samples were collected from Friesian-Holstein calves at the ages specified in chapter 4. Sheep blood and tissue samples were collected from two Blackface ewes aged 6 months. All animals were in normal health at time of sampling. Ovine lymph samples were taken from 10-month old Blackface-cross ewes during the course of an experiment to monitor the effect of *Telodorsagia circumcincta* infection on the abomasum (see section 2.1.2).

2.1.1 Post-mortem and collection of tissue samples

Ovine and calf tissue samples (from thymus, peripheral lymph node, mesenteric lymph node, spleen, and small and large intestinal walls) were collected immediately post-mortem following euthanasia by captive-bolt and pithing. Tissue samples were obtained aseptically, using clean gloves and sterile cutting instruments, to avoid potential contamination. Slices of tissue <5mm thickness were prepared and collected in RNAltr (Applied Biosystems, Warrington, UK). Samples were stored in RNAltr at 4°C until further processing (within 24hours), and stored longer-term in RNAltr at -20°C.

2.1.2 Cannulation of the efferent gastric lymph duct

Efferent gastric lymph samples were collected from sheep as part of an experiment conducted by Dr. W. D. Smith, Moredun Research Institute, Roslin, to study local responses to *Telodorsagia circumcincta* infestation of the abomasum. Experimental sheep had undergone surgical implantation of a cannula into the common gastric

lymph duct, from which lymph samples were collected daily into sterile storage tubes (Smith et al. 1981). Total RNA was provided from these and stored at -80°C.

2.2 Chemicals, solutions and plastic ware

All laboratory plastic ware used in preparation and analysis of nucleic acid was sterile, DNase/RNase-free and disposable. Laboratory glass and plastic ware were sterilised by autoclaving under pressure at 124°C for 15 minutes. Stock solutions and buffers were prepared using either MilliQ ion-exchange purified water (Millipore Milli-RO60 Plus, Millipore, Watford, UK) or sterilised distilled water. Laboratory chemicals used were of analytical or molecular biology grade and were obtained from Sigma-Aldrich (Sigma-Aldrich Co. Ltd., Dorset, UK) unless otherwise stated.

2.3 Cellular techniques

2.3.1 Cell culture

2.3.1.1 Isolation of Peripheral Blood Mononuclear Cells from blood

Blood was collected by jugular venepuncture into an equal volume of Alsever's solution (Appendix A 8.1.1), which served as an anti-coagulant and diluent. Peripheral blood mononuclear cells (PBMC) were harvested using a published protocol (Boyum 1968; Goddeeris and Morrison 1988). Aliquots of 30ml blood/Alsever's mixture were carefully overlaid onto 20ml Ficoll-Paque Plus (GE Healthcare Life Sciences, Buckinghamshire, UK) and centrifuged at 900 x *g* for 30 minutes. The visible layer, containing PBMC at the Alsever's/Ficoll interface, was transferred to a fresh tube and centrifuged at 450 x *g* for 10 minutes; the resulting cell pellet was washed three times using Alsever's solution. Following the final

wash, the cell pellet was resuspended in 5 or 10ml of Alsever's solution to facilitate counting.

In cases where excessive contamination of the cell pellet by red blood cells (rbc) was apparent, rbc-lysis was performed during the pellet-washing process. The cell pellet was resuspended in 2-5ml of rbc-lysis buffer (Appendix A 8.1.2) pre-heated to 37°C, and incubated for 1-5 minutes at 37°C before the washing process was continued as above. PBMC not for immediate use were centrifuged to pellet the cells and excess supernatant poured off before the cell pellets were resuspended and cells were lysed by adding 1ml Tri-Reagent per 5-10 x 10^6 cells and repeat pipetting. Lysed cells were stored at -70°C for up to 1 month.

2.3.1.2 Cell counts

The cell-suspension was diluted 1:10 with 0.4% trypan blue stain (Gibco, Paisley, UK). The cell chamber of an improved Neubauer haemocytometer was filled with the diluted sample and examined under a microscope at 100x magnification. The cell counts within the central square and four corner squares were recorded, and their mean determined. The concentration of cells/ml was calculated with the following formula:

$$\text{cells/ml} = \text{mean no. of cells per square} \times \text{dilution factor} \times 10^4$$

2.3.1.3 Generation and maintenance of hybridomas

Hybridoma cell-lines producing IL-A12, IL-A51, IL-A105, and CC15 (see section 2.3.3.1, table 2-1), as well as murine thymocytes, were retrieved from liquid nitrogen storage. Samples were incubated in a 37°C water bath, until just defrosted, and transferred into 10ml of standard culture medium (SCM) (Appendix A 8.1.3), before

centrifugation at 450 x *g* for 5 minutes to pellet the cells. The supernatant was removed and hybridoma cells were diluted in 2ml alternative SCM containing 20% foetal bovine serum, while thymocytes were diluted in 2ml SCM. Samples were plated out in 24-well plates at a density of 2 x 10^6 cells/ml, incubated in humidified 5% CO_2 at 37°C, and monitored for growth/infection over 36 hours. After this period, thymocytes were added at a density of 1 x 10^5 cells/ml to the hybridoma samples. Cells were maintained in the log growth phase by regular passage, splitting 1:5 (removal of 80% of medium and cells, and replacement with fresh SCM) just prior to confluence. 100mls of supernatant were harvested from each hybridoma cell type by centrifuging the cultured cell suspensions at 450 x *g* for 5 minutes to pellet cells, allowing the removal of the supernatant, which was stored at -20°C.

Supernatants harvested from hybridoma cells were assessed for presence of antibodies by flow cytometry analysis (as described in section 2.3.3.2) of bovine PBMC. IL-A12, IL-A51, IL-A105 and CC15 supernatants were used for primary antibody staining of PBMC at serial two-fold dilutions from neat to 1 in 256. Secondary staining was with a fluorescein isothiocyanate (FITC)-labeled anti-mouse IgG antibody (Invitrogen Ltd., Renfrew, UK). Final dilutions of 1:5 or 1:10 were used for subsequent flow cytometry analyses.

2.3.1.4 Generation and maintenance of *Theileria*-transformed cell-lines

Theileria-transformed T-cell, B-cell, and macrophage cell-lines were provided by Dr. Niall MacHugh, University of Edinburgh, and had been generated and maintained according to published protocols (Goddeeris and Morrison 1988). Cell lines used were: 592TPM, a B-cell line infected with *T. parva*; 663TA/B, a B-cell line infected with *T. annulata*; 592TAA and 663TA, macrophage cell lines infected with *T. annulata*; 109TPM, a CD8 T-cell line, 158TPM, a T-cell line of mixed CD8/γδ phenotype, 951TPM, a γδ T-cell line – all T-cell lines were infected with *T. parva*. RNA was extracted from harvested cells using Tri-Reagent (Sigma-Aldrich Co. Ltd., Dorset, UK) according to manufacturer's instructions (see section 2.4.1.1).

2.3.2 Gene transfer

2.3.2.1 Cos-7 cells and cell passage

The monkey fibroblast cell line Cos-7 (provided by Dr. Jane Hart, University of Edinburgh) was used for transient expression of the bovine MR1 gene and a MR1-chimaeric construct. Cos-7 cells were cultured in DMEM-Cos medium (Appendix A 8.1.4) in T75 cm^2 tissue culture flasks (Costar, Corning Incorporated, Corning, USA) incubated in humidified 5% CO_2 in air at 37°C. Cells were passaged by splitting 1:5 just prior to confluence; medium was removed and the cells were washed with PBS/EDTA (Appendix A 8.1.5), before addition of 5ml 0.25% trypsin-EDTA (1X trypsin solution) (Gibco, Paisley, UK) and incubation at 37°C for 5 minutes to lift off the adherent cells. 10ml DMEM-Cos media was added to dilute the trypsin, and the cell suspension transferred to a 15ml tube and cells pelleted by centrifugation at 450 x *g* for 5 minutes. The cell pellet was resuspended in 5ml, and 1ml transferred back to the original T75 flask along with 18ml fresh DMEM-Cos medium.

2.3.2.2 Transfection

Full-length and chimaeric MR1 and MHC cDNAs cloned into the pcDNA3.1/V5-His-TOPO (Invitrogen Ltd, Renfrew, UK) and pFlag-CMV-3 plasmid vectors (Sigma-Aldrich Co. Ltd., Dorset, UK) (see section 2.4.4) were transfected into Cos-7 cells using FuGENE HD Transfection Reagent (Roche Applied Science, Burgess Hill, UK). Cos-7 cells were plated in a flat-bottomed 96-well plate 24 hours previously at 2 x 10^4 cells in 100µl DMEM-Cos medium per well, and incubated at 37°c in 5% CO_2; Cos-7 cell surface-cover confluence at the time of transfection was approximately 80%. A transfection reaction was initially prepared by warming all reagents to room temperature (RT). For each 10 wells to be transfected, 500µl (serum-free) DMEM and 1µg DNA were combined in a 1.5ml Eppendorf tube. 3µl FuGENE HD was directly added into the solution, and the reaction was incubated at

RT for 15 minutes. DMEM-Cos medium was removed from the pre-prepared 96-well plate, and 50µl of the transfection reaction was added to each well. The plate was further incubated at 37°C, in 5% CO_2 for 4 hours, before the transfection reaction solution was removed and replaced with 100µl DMEM-COS medium, and incubation continued. Cos-7 cells were analysed 48 hours after transfection either as cell suspensions by flow cytometry or as fixed adherent cells by fluorescence microscopy.

2.3.3 Expression analysis

2.3.3.1 Antibodies

Primary antibodies used for flow cytometry analysis were either derived from hybridoma culture supernatants, which were used at a dilution of 1:5 or 1:10 in FACS medium (Appendix A 8.1.6), or from mouse ascitic fluid, which was diluted 1:1000 for use. Primary antibodies used are listed in table 2-1. Secondary antibodies for phenotypic analysis of PBMC were R-Phycoerythrin (R-PE) or FITC-labelled goat anti-mouse immunoglobulin isotype-specific antibodies (IgG1, IgG2a or IgG2b), or labelled anti-mouse IgG antibody, (Invitrogen Ltd., Renfrew, UK / Southern Biotech, Birmingham, Alabama, USA). Secondary antibody used for analysis of MR1 expression in Cos-7 cells was FITC-labelled polyvalent anti-mouse immunoglobulins (G, A, M) antibody (Sigma-Aldrich Co. Ltd., Dorset, UK).

Table 2-1: Bovine monoclonal antibodies used for primary antibody staining for flow cytometry analyses.

Antibody	Specificity	Isotype
MM1A	CD3[a]	IgG1
IL-A12	CD4[b]	IgG2a
IL-A51	CD8 α-chain[c]	IgG1
IL-A105	CD8 α-chain[d]	IgG2a
CC15	WC1 γδ T-cells[e]	IgG2a
GB21A	γδ T-cells[f]	IgG2b
IL-A88	Bovine MHC Ia heavy chain[g]	IgG2a
IL-A19	Bovine MHC Ia in association with β2M[h]	IgG2a
MR1 26.5	Human MR1[i]	ND
W6-32	Human MHC Ia and some bovine MHC Ia alleles[j]	IgG2a
JG-9	Bovine β2M[k]	ND
BG10	Bovine β2M[l]	ND
M2	pFlag-CMV-3 Vector tag[m]	ND

Specifities: a) (Davis et al. 1993) b) (Bensaid and Hadam 1991) c) (MacHugh and Sopp 1991) d) (MacHugh and Sopp 1991) e) (Morrison and Davis 1991) f) (Davis et al. 1996) g) (Toye et al. 1990) h) (Bensaid et al. 1989) i) (Huang et al. 2009) j) (Maziarz et al. 1986) k) & l) were provided by Dr. Shirley Ellis, Institute for Animal Health, Compton; their specificities are unpublished m) (Sigma-Aldrich Co. Ltd., Dorset, UK). Isotype ND = not determined; a FITC-labelled anti-mouse immunoglobulins secondary antibody was used for detection of these primary antibodies.

2.3.3.2 Flow cytometry analysis

Flow cytometry was used to identify populations of T-cells in PBMC of cattle and sheep, and to study surface expression of transfected genes on Cos-7 cells. PBMC were isolated from fresh blood samples and counted as in section 2.3.1.1/2, and prepared in a round-bottomed 96-well plate at a concentration of approximately 1 x 10^6 cells/well (in 50ul FACS medium). Cos-7 cells were harvested as described above and were suspended at 2 x 10^7 cells/ml and 50ul added to each well.

Aliquots of 25-50µl of primary monoclonal antibody were added to each well, and the plate incubated at 4°C for 30 minutes, before cells were washed three times in 200µl FACS medium. The same procedure was repeated for secondary antibody. Following the third wash in FACS medium, cells were resuspend in FACS medium and analysed on a BD FACSCalibur Flow Cytometer (BD Biosciences, San Jose, California, USA).

2.3.3.3 Purification of cells by flow cytometry

Cell subsets within bovine PBMC were purified by flow cytometry according to their cell-surface phenotype. Between 10-30mls of isolated PBMC at a density of approximately 1 x 10^7 cells/ml were stained by indirect immunofluorescence with cell subset-specific antibodies as described for cell phenotyping and separated into distinct phenotypic populations using a BD FACSAria Flow Cytometer (BD Biosciences, San Jose, California, USA) and FACSDiva software version 5.02. Once separated, samples were reanalysed by flow cytometry to assess purity.

2.3.3.4 Immunofluorescence microscopy

Cos-7 cells were analysed for gene expression 48 hours after transfection by immunofluorescence staining of adherent cells and microscopic examination with a

Leitz "Laborlux K" microscope (Esselte UK Ltd., Uxbridge, UK) and a Zeiss Axiovert 25s (Carl Zeiss Ltd., Welwyn Garden City, UK) microscope, equipped with epi-fluorescence. Cells were prepared for analysis in two ways:

(i) Cytospins – Transfected cells were harvested from the wells of flat-bottom 96 well plates by treatment with trypsin-EDTA as described above (section 2.3.2.1) and transferred to the wells of a U-bottom 96-well plate. An additional 100µl DMEM-Cos medium was added to each well and after centrifugation of the plate at 450 x *g* for 5 minutes, the medium was flicked off and the cells were resuspended in 200µl DMEM-Cos medium. Double cytofunnel cuvettes (Shandon, Thermo Fisher Scientific Inc., Waltham, MA, USA) were mounted onto superfrost microscope slides in metal holders; 200µl cell suspension was transferred to either side of the cuvette, and the cuvette centrifuged at 800rpm for 3 minutes in a Cytospin3 centrifuge (Shandon, Thermo Fisher Scientific Inc., Waltham, MA, USA). Slides were carefully extracted and allowed to air-dry, before being fixed in an acetone bath for 5 minutes and air-dried again. Cytospin blot perimeters were sealed with Liquid Blocker Super Pap Pen (Daido Sangyo Co. Ltd., Tokyo, Japan). 20-50µl of primary antibody was applied to the cytospin blot, and the slide incubated for 30 minutes in a humidified box, after which it was rinsed twice in a PBS (Appendix A 8.1.7) bath. The same procedure was repeated for secondary antibody. Following the second PBS wash, fluorescent-mounting medium (Dako UK Ltd., Ely, UK) was used to add a cover-slip to the slide, and cytospins were examined by fluorescence microscopy.

(ii) Immunostaining within wells – Medium was removed from the 96 well plate and the cells were washed by addition and removal of 200µl PBS per well. 100µl of an acetone/ethanol mixture was added per well to fix the cells, and the plate was incubated at RT for 5 minutes. Acetone/ethanol was flicked off, and the cells washed with 200µl PBS. 50µl primary antibody was added to each well, and the plate incubated for 30 minutes at 4°C. Following incubation, cells were washed three times in 200µl FACS medium. The same procedure was repeated for secondary antibody. After the final rinse in FACS medium, 50µl PBS was added to each well, and cells in the wells examined by immunofluorescence microscopy.

2.4 Molecular techniques

2.4.1 Preparation of nucleic acids

2.4.1.1 RNA extraction from PBMC

Total RNA was extracted from cell pellets of PBMC using Tri-Reagent (Sigma-Aldrich Co. Ltd., Dorset, UK). Cell pellets were resuspended in 100µl of Alsever's solution in Eppendorf tubes, and cells were lysed by adding 1ml Tri-Reagent per 5-10 x 10^6 cells and repeat pipetting, followed by incubation at RT for 5 minutes. Chloroform (0.2ml per 1ml of Tri-Reagent) was added to the solution, vortexed for 15 seconds, and incubated at RT for 15 minutes. The solution was centrifuged at 12,000 x *g* for 15minutes at 4°C to separate it into three phases – an organic (protein) phase, an interphase (DNA), and an aqueous phase (RNA). The RNA phase was collected and isopropanol (0.5ml per 1ml Tri-Reagent) added, incubated at RT for 10 minutes, and centrifuged at 12,000 x *g* for 10 minutes at 4°C to precipitate the RNA. The RNA pellet was washed twice using 1ml 75% ethanol and allowed to air-dry. Once the ethanol had evaporated, the pellet was resuspended in 25µl of nuclease-free water. RNA was stored at -70°C.

2.4.1.2 RNA extraction from tissue

(i) Homogenisation of tissue: 100mg of (thymus, spleen, peripheral lymph node, mesenteric lymph node, small intestinal wall, ileal wall) tissue was weighed and diced into small pieces using autoclaved, RNA-zapped instruments. This tissue was placed in a FastRNA ProGreen matrix tube (MP Biomedicals, Solon, Ohio, United States) containing 1ml RLT-β-Mercaptoethanol (100µl 2-Mercaptoethanol/1ml RLT) (Sigma-Aldrich Co. Ltd., Dorset, UK). The sample was processed in the FastPrep instrument (MP Biomedicals) in runs of 40 seconds at a speed setting of 6.0, applying 1-3 runs of processing, depending upon the tissue-sample, and a cooling

period of 10 minutes between each run. The resulting liquid upper layer in the matrix tube was transferred to a Qiashredder tube (Qiagen, Crawley, UK) and centrifuged to filter the RNA lysate and remove any remaining debris.

(ii) Extraction of RNA: RNA extraction from tissue samples was done using the RNeasy mini-kit (Qiagen, Crawley, UK). 1400µl of RNA lysate was mixed with an equal volume of 70% ethanol and run through an RNeasy column, which was centrifuged to trap the RNA in the RNeasy filter. The trapped RNA was then washed by passing 700µl of RW1 buffer through the column, followed by two washes with 450µl RPE/ethanol buffer. The column was transferred to a "no stick" Ambion eppendorf tube, and the RNA eluted by adding 50µl nuclease-free water to the column, and centrifugation.

2.4.1.3 Preparation of plasmid DNA

(i) Mini-preps: *E. coli* competent cells containing transformed DNA were expanded in 6ml of Luria-Bertani (LB) medium by incubation overnight at 37°C while shaking at 220rpm. Plasmid DNA was recovered using either Wizard Plus SV MiniPreps DNA Purification System or Pure Yield Plasmid MiniPrep System (Promega, Hampshire, UK). 3ml of the expansion culture was pelleted by centrifugation at 10,000 x g for 5 minutes and the supernatant removed. 200µl cell resuspension solution, 200µl cell lysis solution, and 200µl neutralisation solution (all provided with the Promega kit) were added to the pelleted bacteria, the contents were mixed and then centrifuged to pellet waste products. The clear supernatant was added to 1ml DNA resin and passed through a mini-column to allow DNA to bind to the mini-column resin. 2ml of column wash solution was passed through the mini-column, and the resin dried by centrifugation. 50µl of nuclease-free water was added to the resin, and the mini-column centrifuged to elute the DNA.

(ii) Maxi-preps: *E. coli* competent cells containing transformed DNA were expanded for 5 hours in 6ml of LB medium, then overnight in 200ml LB medium in

a wide-based beaker. Plasmid DNA was recovered using the Qiagen Endofree Plasmid Maxi Kit (Qiagen, Crawley, UK) as per the manufacturer's instructions. Briefly, 200ml of the expanded bacterial culture was pelleted by centrifugation at 6000 x g for 15 minutes at 4°C, and the supernatant removed. Pellets were resuspended in 10ml solution P1, then 10ml solution P2 was added, and the reaction mixed and incubated at RT for 5 minutes, before 10ml solution P3 was added and vigorously mixed. The reaction was incubated at RT for 10 minutes, and insoluble proteins were filtered out by passage through a QIAfilter Maxi Cartridge. 2.5ml ER buffer was added to the filtered lysate, mixed, and incubated on ice for 30 minutes. During this incubation a Qiagen-tip 500 column was prepared by filtration of 10ml buffer QBT through the column matrix. The lysate/ER was then added to the column and DNA allowed to bind to the column resin by gravity flow. Two washes of the column with 30ml buffer QC were performed, and the DNA eluted by allowing 15ml buffer QN to flow through the resin. The DNA was precipitated by centrifugation with 10.5ml 100% isopropanol at 15000 x g for 30 minutes at 4°C and then 5ml of endotoxin-free 70% ethanol at 15000 x g for 10 minutes at 4°C. The DNA pellet was then air-dried and dissolved into 100µl TE.

2.4.1.4 DNase treatment of RNA for qPCR

RNA from PBMC and tissue samples to be analysed by qPCR was treated with RQ1 RNase-free DNase (Promega, Hampshire, UK) as per the manufacturer's instructions. 2µg RNA were mixed with 1µl RQ1 10X Reaction Buffer, 2µl RQ1 DNase (1u/µl), and made up to 10µl with nuclease-free water. This mixture was incubated at 37°C for 30 minutes in a hot-lid thermal-cycler, before 1µl RQ1 DNase Stop Solution was added, and the reaction incubated at 65°C for 10 minutes. Final RNA concentration for reverse transcription was 1µg in 5.5µl.

2.4.1.5 Restriction endonuclease digestion of DNA for cloning

DNA was cloned into the pFlag-CMV-3 expression vector (Sigma-Aldrich Co. Ltd., Dorset, UK) through restriction endonuclease digestion of the vector and appropriate inserts. Vector restriction endonuclease sites not present in the insert sequence were selected; these were added to either end of the sequences to be inserted by incorporation into the PCR primers used to amplify the sequence, each flanked by six additional random nucleotides to aid the digestion reaction. The pFlag-CMV-3 vector and full-length MR1 cDNA insert were digested with HindIII (New England Biolabs (NEB) Ltd., Hitchin, Herts., UK) and XbaI (NEB) as per the manufacturer's guidelines. Briefly, 500ng DNA was prepared with 5µl 10X NEB Buffer 2, 0.5µl Bovine Serum Albumin (BSA), 1µl HindIII, 1µl XbaI, and made up to 50µl with nuclease-free water. This reaction was incubated at 37°C for 3 hours, then 65°C for 20 minutes. The components of the MR1 chimaeric insert were digested with HindIII and BsiWI (NEB) for the MR1 component, and XbaI and BsiWI for the MHC component. Reactions were set-up with 500ng insert DNA, 5µl 10X NEB Buffer 2, 1µl HindIII (cMR1) or 1µl XbaI and 0.5µl BSA (cMHC), and nuclease-free water to a volume of 50µl This mixture was incubated at 37°C for 3 hours, then 65°C for 20 minutes. 1µl BsiWI was then added to the reaction mixture, and further incubations at 55°C for 3 hours then 80°C for 20 minutes performed.

Digested DNA was purified using the QIAQuick PCR Purification Kit (Qiagen, Crawley, UK) as per the manufacturer's guidelines – 250µl buffer PB was added to the reaction mixture and passed through a QIAQuick spin column to allow DNA fragments over 80bp to attach to the column membrane. The membrane was washed in buffer PE, before being allowed to dry, and DNA eluted by passage of 30µl of nuclease-free water across the column membrane. Product DNA was quantified and cloned as in section 2.4.4.

2.4.2 Polymerase chain reactions

2.4.2.1 Reverse Transcription (RT)-PCR for cDNA synthesis

cDNA was synthesised from RNA using the Reverse Transcription System (Promega, Hampshire, UK). 1µg of RNA was incubated at 70°C for 10 minutes. A reaction mixture of 4µl 25mM $MgCl_2$, 2µl Reverse Transcription 10X buffer, 2µl 10mM dNTP mixture, 0.5µl recombinant RNAsin ribonuclease inhibitor, 15u avian myeloblastoma virus reverse transcriptase (AMV-RT), 0.5µg oligo$(dT)_{15}$ primer, was created, the 1µg of RNA added to this, and the volume made to 20µl with nuclease-free water. This reaction was incubated at 42°C for 1 hour to create cDNA, heated to 95°C for 5 minutes, then incubated at 0°C for 10 minutes to inactivate the AMV-RT. cDNA was stored at -20°C.

2.4.2.2 Amplification of DNA by PCR

Working stock primers were used at 10pmol/µl. Unless stated otherwise, PCR reactions were run in a total volume of 20µl (or multiples thereof) composed of 10pmol forward primer, 10pmol reverse primer, 2µl ABgene custom PCR master mix SM-0005 (Appendix A 8.1.8) (ABgene, Epsom, Surrey, UK), 0.2µl (5units/µl) BIOTAQ DNA polymerase (Bioline, London, UK), 1µl cDNA, and made up to 20ul with sterile water. Hot-lid thermal-cyclers were used, either PTC-100 (MJ Research Inc, Watertown, Mass, USA) or G-Storm (GRI Ltd., Essex, UK) machines. The standard cycling protocol for PCR was incubation at 95°C for 10 minutes, 35 cycles of (95°C for 1 minute, specific annealing temperature for 1 minute, 72°C for 1 minute), and 72°C for 10 minutes, except where specified. Specific annealing temperatures are listed within each chapter.

2.4.2.3 RACE PCR

RACE (5' and 3' Rapid Amplification of cDNA Ends) PCR was performed using the SMART RACE cDNA Amplification Kit (Clontech Laboratories, Inc.). SMART (Switching Mechanism at 5' end of RNA Transcript) first-strand cDNA synthesis allows generation of RACE-ready full-length cDNA through reverse transcription. The dC-tailing action of Moloney murine leukaemia virus reverse transcriptase (MMLV-RT) through terminal transferase activity allows annealing of the SMART II A oligonucleotide to the original RNA sequence, which serves as an extended template for creation of cDNA through further reverse transcription. RACE-ready cDNA is subsequently used directly in RACE PCR reactions, in combination with gene-specific primers and universal primer mix (UPM, which contains two primers, short and long universal primers (SUP/LUP)) which suppress inverted repeat elements and allow gene-specific DNA synthesis. The resulting RACE products were purified, cloned and sequenced to create a library of gene transcripts.

(i) First strand cDNA synthesis: was performed as per the Clonetech Laboratories, Inc. protocol. Briefly, 1µg of RNA was mixed with 1µl 5'-CDS primer A and 1µl SMART II A oligo, and made up to 5µl with sterile water. This preparation was incubated at 70°C for 2 minutes, then incubated on ice, before adding 2µl 5X First-Strand Buffer, 1µl 20mM DTT, 1µl 10mM dNTP mix and 1µl Primescript MMLV Reverse Transcriptase (Takara Bio Inc, Japan). The reaction was mixed and incubated for 90 minutes at 42°C in a PTC-100 thermal cycler. Reaction product was diluted by adding 100µl tricine-EDTA buffer (10mM tricine-KOH at pH8.5; 1.0mM EDTA) and incubated at 72°C for 7 minutes. cDNA produced was stored at -20°C.

(ii) Rapid Amplification of cDNA Ends: cDNA was amplified as per the Clonetech protocol. Briefly, reaction master mix was created using 34.5µl sterile water, 5µl 10X Advantage 2 PCR buffer, 1µl 10mM dNTP mix, 1µl 50 Advantage 2 Polymerase mix, and mixed well. 2.5µl 5'-RACE-ready cDNA, 5µl 10X UPM, and 1µl 10µM bovine TCRα constant region primer BovC4r were mixed in a separate tube, and the master mix added giving a total volume of 50µl. Two negative controls

were also run – one without 10X UPM, and one without BovC4r. Reactions were incubated in a PTC-100 thermal cycler for 5 cycles of (30 seconds at 94°C, 3 minutes at 72°C), 5 cycles of (30 seconds at 94°C, 30 seconds at 70°C, 3 minutes at 72°C), and 25 cycles of (30 seconds at 94°C, 30 seconds at 68°C, 3 minutes at 72°C).

2.4.2.4 'Real-time' quantitative PCR

'Real-time' quantitative PCR was used to compare expression of the MAIT TCRVα gene to that of the TCRα constant region, identifying the relative abundance of MAIT cells within the overall αβ T-cell population. RNA was isolated from ovine and bovine tissues using RNeasy Protect kit (see section 2.4.1.2) and DNase-treated using Promega RQ1 RNase-free DNase (section 2.4.1.4). cDNA synthesis was performed using AMV reverse transcriptase, as in section 2.4.2.1. The qPCR used primers (bovine: "I"/"Jr"; ovine: "boovUntr3f"/"CorovC3r"), designed to anneal to the TCRα V segments that are used by bovine/ovine MAIT cells and the respective α constant chains, yielding 750bp PCR products (for primer sequences, see section 3.2.3). These were sequenced to confirm MAIT TCRα identity and used as positive controls and to generate qPCR standard curves.

For the MAIT-cell V-region-specific qPCR the following primers/probe were used – forward (both species): "boovAV19f"; reverse (bovine): "boovAJ33r"; reverse (ovine): "ovAJ33r"; probe (both species): "boAV19". For the α chain constant-region-specific qPCR, primers "NGconst-f" and "NGconst-r", and probe "NGconst" were used for both cattle and sheep. Primer sequences are provided in chapter four. Primers and probes were made by Applied Biosystems, Warrington, UK. qPCRs were performed in a 20µl reaction volume containing 500nM of each primer, 250nM of probe, and 3µg of cDNA in Taqman Gene Expression Master Mix (Applied Biosystems, Warrington, UK). Reactions were run in an ABI7500 qPCR machine (Applied Biosystems, Warrington, UK) under the following cycling conditions: an initial incubation of 2 minutes at 50°C, an incubation for 10 minutes at 95°C, and 50 cycles of 10 seconds (ovine 15 seconds) at 95°C and 1 minute at 53°C. All standards

and samples were run in triplicate, and results expressed as the mean value with standard error provided. Results were analysed using Applied Biosystems 7500 Fast System SDS software, version 1.4 (Applied Biosystems, Warrington, UK).

In chapter four, qPCR results for two 3-month old calves were provided by Dr. Ildiko Van Rhijn, University Utrecht, Netherlands. Specific method for these results is described in chapter 4.

2.4.2.5 Primer design and production

Standard and RACE PCR primers were designed either manually or using DNASIS software (see 2.5.2) and ordered from Eurofins MWG Operon (Ebersberg, Germany).

2.4.3 Analysis of nucleic acids

2.4.3.1 Quantification of RNA and DNA

Quantification of RNA and DNA was performed using either the Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA), or the GeneQuant spectrophotometer (GE Healthcare Life Sciences, Buckinghamshire, UK). Samples were diluted with nuclease-free water, and absorbance readings taken at 260nm/280nm wavelengths. Concentrations were assessed using the following formulas:

RNA (ng/µl) = 260nm reading x dilution factor x 40

DNA (ng/µl) = 260nm reading x dilution factor x 50

2.4.3.2 Housekeeping genes

GAPDH and G3PDH were used as housekeeping genes for ovine and bovine cDNA, respectively. Primers used to detect these were "ovGAPDHf" (5'-AAG GCA GAG AAC GGG AAG-3') and "ovGAPDHr" (5'-ACT GTC CAC GCC ATC ACT-3'), used at annealing temperature (Tm) 56°C; and "G3PDH1" (5'-GAT GCT GGT GCT GAG TAT GTA GTG-3') and "G3PDH2" (5'-CTC CCA ACG TGT CTG TTG TGG AT-3') at Tm: 60°C.

2.4.3.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to study the molecular weight of PCR products as a guide to reaction success and PCR product identity.

(i) 1.5% gel solution of agarose (Bioline, London, UK) in 1 x TAE buffer (Appendix A 8.1.9) was used to create gels for standard analysis in a midi-gel system (Flowgen, Ashby de la Zouch, Leics., UK). PCR products were mixed 2:1 with loading buffer (Appendix A 8.1.10) and loaded into a gel placed in 1 x TAE buffer, through which an 80-120V charge was run for 30-120 minutes, as appropriate. 1µg of 1Kb Plus DNA Ladder (Invitrogen Ltd., Renfrew, UK) was used to calibrate the sizes of different PCR products. DNA staining was either performed using Biotium GelRed 10,000X (Cambridge BioScience, Cambridge, UK) at 1:10,000 dilution within the agarose gel solution, or by staining the gel in an ethidium bromide bath (0.5µg/ml in 1 x TAE buffer) for 30 minutes following completion of electrophoresis. Stained gels were visualised using a BioRad Gel Doc 2000 or BioRad Molecular Imager FX Pro Plus, and BioRad Quantity One software (Biorad, Hercules, CA, USA).

(ii) 3% Metaphor gel solution was used to differentiate small, similarly-sized fragments of DNA. MetaPhor agarose (Lonza, Rockland, ME, USA) in 1 x TAE buffer was used to create gels for analysis. Gel Red was added to the gel solution to

stain DNA product. PCR products were mixed 2:1 with loading buffer and loaded into a gel placed in 1 x TAE buffer, through which a 120V charge was run for 3.5 hours. 1.5µg 50bp DNA Step Ladder (Promega, Hampshire, UK) was used to estimate molecular weight of products. Gels were visualised using a BioRad Molecular Imager FX Pro Plus and Quantity One software.

2.4.3.4 DNA extraction from agarose

DNA samples were extracted from agarose gels using either a Geneclean II kit (Qbiogene, Cambridge, UK), Qiagen MinElute Gel Extraction kit (Qiagen, Crawley, UK), or QIAQuick Gel Extraction kit (Qiagen), as per the manufacturers' instructions. DNA bands were excised from agarose gel using sterile instruments and weighed to estimate volume (0.1g = 100µl). The Geneclean II kit was used for the vast majority of reactions - three times the gel volume of chaotropic salt NaI was added to keep the Na concentration above 4M (thus forcing the DNA into the silicate). This solution was incubated at 55°C in a water bath until the gel had dissolved. 15µl of glassmilk (silicate) was added and the solution incubated at RT for 10 minutes to allow DNA binding to the silicate. The DNA/silicate was pelleted by centrifugation and washed three times in 700µl new wash solution. The supernatant was removed and the pellet allowed to air dry thoroughly. DNA was eluted by resuspending the pellet in 15µl nuclease-free water – this lowers the Na concentration releasing the DNA into the supernatant. The solution was centrifuged and the supernatant collected; this process was repeated twice to remove as much silicate as possible. Supernatant containing DNA was stored at -20°C.

2.4.3.5 PCR product purification

PCR products were purified for cloning by removing primers and salts, using the Wizard PCR Preps DNA Purification System (Promega, Hampshire, UK). 100µl of

direct purification buffer and approximately 100µl of PCR product were mixed, and 1ml of DNA resin added before vortexing the solution vigorously three times within one minute. The resulting homogenate was passed through a Wizard PCR Preps DNA Purification System mini-column to adhere the product DNA to the mini-column resin. The resin was subsequently rinsed with 2ml of 80% isopropanol and dried by centrifugation. 40µl of nuclease-free water was used to elute the DNA from the resin by centrifugation.

2.4.3.6 DNA Sequencing

Sequencing of small numbers of DNA samples was performed by DBS Genomics at Durham University, UK, using an Applied Biosystems 3730 DNA Analyser (Applied Biosystems, Warrington, UK). Samples were either purified PCR product (as described in section 2.4.3.5) with appropriate primers, or purified plasmid DNA (see section 2.4.1.3) with specific or vector primers as appropriate.

2.4.3.7 96-well plate sequencing reactions

Colonies transformed with plasmid vector containing product from RACE-PCRs were sequenced in a 96-well plate format, using DNA extraction methods and sequencing reaction provided by Dr. T. Connelley, University of Edinburgh, which was based upon a BigDye Terminator v3.1 (Applied Biosystems, Warrington, UK) protocol adapted from that provided by the Sequencing Unit of the Department of Zoology, University of Oxford.

Sample Preparation: Colonies were transferred into wells containing 20µl nuclease-free water within a 96-well PCR plate, filling 90 wells per plate and leaving 6 wells as negative controls. Plates were sealed and incubated at 95°C for 10 minutes in a hot-lid thermal-cycler to denature DNA, before centrifugation at 600 x *g* for 5 minutes to pellet bacterial debris. Supernatant containing DNA was transferred to a new 96-well plate. To purify the DNA, 60µl of 20% polyethylene glycol (PEG)

MW8000/2.5M NaCl (Appendix A 8.1.11) was added to each well, and the plate vortexed and incubated either overnight at 4°C or at RT for 30 minutes. The plate was sealed and centrifuged at 2250 x g for 1 hour at 4°C to pellet (larger) DNA. Supernatant was removed by centrifugation of the inverted plate at 500 x g for 1 minute. Two washes with 150µl ice-cold 70% ethanol were performed, before the pellet was air-dried and resuspended in 50µl sterile water. Samples were stored at -20°C. At this stage, a sample from each well was mixed with loading buffer and run on a 1.5% agarose gel to confirm the presence of product.

Sequencing Reaction: Each reaction was composed of a master mix of 4µl of 0.67pmol/µl sequencing primer BovCintr (5'- TGC TCA ACA TGG AGA TCC TG-3'), 2µl 5X CSA buffer (Appendix A 8.1.12), 0.25µl BigDye Terminator v3.1, and 1.75µl nuclease-free water, to which 2µl of purified PCR product was added. The reactions were prepared in 96-well plate format, and the plate sealed and incubated for 30 cycles at (96°C x 10 seconds, 50°C x 5 seconds, 60°C for 2 minutes). 15µl nuclease-free water, 50µl absolute ethanol and 2µl of 3M NaOAc pH5.2 (Appendix A 8.1.13) were added to each well, the plates sealed and vortexed, and incubated at RT for 45 minutes. Sequencing products were pelleted by centrifugation at 2250 x g for 75 minutes at 4°C, and supernatant discarded as described earlier. DNA pellets were washed once in 150µl ice-cold 70% ethanol before air-drying. Prepared 96-well plates were sent to the Sequencing Unit, Department of Zoology, University of Oxford for analysis using an Applied Biosystems 3730xl DNA Analyser (Applied Biosystems, Warrington, UK).

2.4.4 Cloning into plasmid vectors

2.4.4.1 Vectors used

(i) Standard sub-cloning of PCR products was performed using the pGEM-T Easy System (Promega, Hampshire, UK), which contains a 3015bp vector (Appendix A 8.2.1).

(ii) Cloning of MR1 and a MR1-chimaera was performed into a 5'-tagged 6kb vector, pFlag-CMV-3 expression vector (Sigma-Aldrich Co. Ltd, Dorset, UK) (Appendix A 8.2.3). The original vector was transformed directly into JM109 cells to allow expansion and extraction (see section 2.4.1.3) of vector DNA.

(iii) Plasmid DNA from a bovine classical MHC class I (N*02401) transformed into a TOPO TA 5.5kb expression vector (Appendix A 8.2.2) using the Invitrogen pcDNA3.1 / V5-His TOPO TA Expression Kit (Invitrogen Ltd, Renfrew, UK), was provided by Dr. Jane Hart, University of Edinburgh.

2.4.4.2 Ligation and cloning

(i) Purified PCR product was ligated into pGEM-T Easy vector using a reaction mixture of 5µl 2X rapid ligation buffer, 1µl pGEM-T Easy vector, 3µl PCR product, 1µl T4 DNA ligase (3 weiss units/µl). This reaction mixture was incubated either for 1 hour at RT or overnight at 4°C.

(ii) Vector DNA and insert DNA were prepared by restriction endonuclease digestion (see section 2.4.1.5). A ligation reaction of insert DNA and vector DNA at a ratio of 4 : 1 was mixed with appropriate volumes of T4 DNA ligase (to provide 3 Weiss units per 10µl total reaction volume) and 2X rapid ligation buffer (to provide 50% of total reaction volume). The reaction mixture was incubated for 1 hour at RT then overnight at 4°C.

2.4.5 Transformation and analysis of transformed *E. coli*

2.4.5.1 Transformation and competent cells

(i) PCR products in pGEM-T Easy vector were transformed into JM109 High Efficiency Competent Cells (Promega, Hampshire, UK). 50µl of JM109 cells were carefully added to 5µl of the ligation product, and incubated at 4°C for 1 hour, before being heat-shocked for 45 seconds in a 42°C water bath and returned to 4°C for a further two minutes. 950µl of Super Optimal broth with Catabolite repression (SOC) medium (Appendix A 8.1.14) was added and the solution incubated at 37°C for 1 hour while shaking at 150rpm. A prepared Luria-Bertani (LB) agar medium plate containing 100µg/ml ampicillin (Sigma-Aldrich Co. Ltd, Dorset, UK) (Appendix A 8.1.15) was coated with a mixture of 100µl of 100mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Bioline, London, UK) and 20µl of 50mg/ml 5-bromo-4-chloro-3-indolyl-b-D-galactosidase (X-Gal) (Bioline, London, UK), and dried. 100-120µl of the transformation solution was spread onto the agar plate and incubated at 37°C overnight to allow colony growth.

(ii) Full-length MR1 and MR1 chimaera in pFlag-CMV-3 vector were transformed into JM109 as in section 2.4.5.1(i), except that IPTG/X-Gal was omitted from the agar plate.

2.4.5.2 Selection and growth of antibiotic-resistant colonies

(i) Antibiotic-resistant colonies were selected on the basis of colour, where white colonies indicated that the expression of the pGEM-T Easy vector Lac Z gene had been interrupted by the insertion of a DNA fragment, thus preventing β-galactosidase formation and hydrolysation of X-Gal to a blue colour. Alternately, blue colonies represented Lac Z gene expression

indicating absence of insert. White colonies were tested for insert size by PCR with either gene-specific primers using appropriate parameters, or vector primers T7 and SP6 using an annealing temperature of 55°C and standard cycling protocol (see section 2.4.2.2).

(ii) Antibiotic-resistant pFlag-CMV-3 vector colonies were tested for insert size by PCR with vector primers N-CMV-30 (5'-AAT-GTC-GTA-ATA-ACC-CCG-CCC-CGT-TGA-CGC-3') and C-CMV-24 (5'-TAT-TAG-GAC-AAG-GCT-GGT-GGG-CAC-3') using an annealing temperature of 55°C and standard cycling protocol (section 2.4.2.2).

In each case, positive colonies were selected and transferred to a numbered grid on a second agar plate, and incubated for 6-18 hours at 37°C. Colonies to be amplified for extraction of plasmid DNA were grown in LB medium with 50µg/ml ampicillin (Sigma-Aldrich Co. Ltd, Dorset, UK) (Appendix A 8.1.16) by overnight incubation at 37°C while shaking at 224rpm.

2.4.5.3 Preparation of glycerol stocks

To create a frozen stock of each plasmid, 150µl of sterile glycerol was added to 750µl of the overnight expansion culture, vortexed briefly, and immediately stored at -70°C.

2.5 Sequence data and analysis

2.5.1 Sequence databases & search engines

Original sequence data and published gene data were obtained using the online sequence databases of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/), Ensembl, a joint operation between the European

Bioinformatics Institute (EMBL-EBI) and the Wellcome Trust Sanger Institute (http://www.ensembl.org/index.html), and 'IMGT', the international ImMunoGeneTics information system (http://imgt.cines.fr/). Within these sites, the Basic Local Alignment Search Tool (BLAST) was used to detect matches to known sequence within the published human, murine, bovine and ovine gene databases.

2.5.2 Sequence analysis

The computer software DNASIS MAX, Version 2.07.000.001, 2006 (Hitachi Software Engineering Company, Ltd.) was used for analysis and study of gene sequences and sequencing data. Internal databases were established using DNASIS software to allow BLAST analysis of multiple sequences. The DNASIS primer design functions were also used for designing the majority of PCR primers used.

2.5.3 TCR sequence nomenclature

Current TCRVα nomenclature is designed to follow the genomic order of the genes on the TCRα locus. However, it is not yet possible to apply this nomenclature to cattle due to the incomplete assembly of the genome, including the bovine TCRαδ locus. Therefore to avoid confusion, and in conjunction with all current literature on this topic, the previous nomenclature system defined by WHO-IUIS and based on the work by Arden et al (Arden et al. 1995), is used for all species throughout this study. In this system, V genes were numbered in the order in which they were discovered; orthologous V genes in the different species therefore have different numbers.

3 Molecular Cloning and Characteristics of Ruminant MAIT

3.1 Introduction

Mucosal associated invariant T (MAIT) cells are characterised by the expression of an invariant TCRα chain. In humans and mice these chains are encoded by the Vα7.2-Jα33 and Vα19-Jα33 gene segments respectively. These sequences show a high level of cross-species conservation, suggesting an evolutionarily-conserved function for these cells. Previous studies also demonstrated a homologous bovine partial TCRα chain sequence, again displaying marked conservation of sequence with human and mouse (Tilloy et al. 1999). This sequence had been identified in an earlier study during analyses of randomly selected bovine TCRα chains (Ishiguro et al. 1990). Apart from these observations, all of the work on characterisation of MAIT cells has been done in mice and humans. MAIT cells are more abundant in humans than in mice, representing up to 4% of circulating T-cells, and up to 15% - 40% of the double-negative (DN) $CD4^-/CD8^-$ T-cell population; in mice they represent <2% of $CD4^-/CD8^-$ T-cells (Tilloy et al. 1999; Treiner et al. 2005; Martin et al. 2009).

The adaptive immune system relies upon a broad antigen-receptor repertoire to rapidly recognise any of the potential pathogens with which it may be presented. Accordingly, the size and diversity of an animal's TCR repertoire have been linked to efficient protection against pathogens and disease (recent reviews by Nikolich-Zugich et al. 2004; Turner et al. 2006). A number of genetic features contribute to the ability to produce a diverse TCR repertoire during the normal genesis of T-cells. Thus, the functional gene encoding each of the two polypeptide chains of the TCR is generated during T cell development by rearrangement of V, D (β chain only) and J segments, drawn from a genomic library of these gene segments. Further diversity is created by N-region junctional addition and deletion of nucleotides. This allows the generation of a very large repertoire of clones with distinct receptors. Since these clones undergo little, if any, peripheral expansion in the absence of antigen, they are each present at a very low frequency within the naïve T cell repertoire. However, the

invariant T-cell subsets, MAIT cells and iNKT cells, differ in this regard in that they are detected at a much higher frequency than expected in the naïve repertoire. Hence, they theoretically can respond more rapidly upon exposure to their respective ligands.

The work described in this chapter set out to characterise MAIT TCRα sequences of cattle and sheep, in order to provide tools that could be used to identify and quantify MAIT cells in these species and to allow further analyses of the conservation of the MAIT TCRα between species. The experiments initially exploited the availability of partial MAIT TCRα sequence data (Tilloy et al. 1999) to generate full-length sequences for this species, which were then used to derive a sequence from sheep. PCR assays designed on the basis of these sequence data were employed to investigate expression of the constituent V and J gene segments, as well as the rearranged MAIT sequence, in bovine PBMC.

3.2 Materials and Methods

3.2.1 Experimental animals

Peripheral blood samples were obtained from healthy, adult Holstein-Friesian cattle and Blackface-cross sheep as described in section 2.1. Peripheral blood mononuclear cells (PBMC) were prepared as in section 2.1.1; these were used for total RNA extraction and cDNA synthesis (see sections 2.4.1.1 and 2.4.2.1) and for cell subset purification and analysis by flow cytometry (section 2.3.3).

3.2.2 Online resources and sequence analysis

Online sequence databases from the National Center for Biotechnology Information (NCBI), Ensembl, and 'IMGT', the international ImMunoGeneTics information system, were used to obtain original sequence data and search for published gene data. Sequence analysis, and comparison to 'in-house' databases, was performed using DNASIS software (Hitachi Software Engineering Company, Ltd.) (see section 2.5). Sequence motifs used for definition of the TCRVα gene CDR3 were based on published definitions (Moss and Bell 1995).

3.2.3 PCR primers and reaction conditions

The PCR primer sequences used are listed in tables 3-1 and 3-2. PCR assays were performed using reaction conditions and reagents as described in section 2.4.2.2. Optimal cycling conditions were assessed for each individual reaction; a standard reaction was then used, consisting of 10 minutes at 95°C, 35 cycles of (1 minute at 95°C, 1 minute at a specific annealing temperature, 1 minute at 72°C), and a final extension period of 10 minutes at 72°C. Optimal annealing temperatures for each primer combination are listed in tables 3-1 and 3-2. PCR products were analysed on 1.5% agarose gels, as described in section 2.4.3.3.

Table 3-1: Primers used in PCR to amplify the bovine and ovine MAIT TCRα sequences, and optimal annealing temperature (Tm) for each reaction. Vα19-Jα33 is taken to represent the ovine and bovine MAIT TCRα variable (V) region. C = TCRα constant region.

Target Sequence	Forward 5' sequence 3'	Reverse 5' sequence 3'	Optimal Tm
Bovine/ovine MAIT TCRα CDR3	boovAV19f CATTCCTTAGACGCTCTGATGCACA	boovAJ33r GCAACTATCAGTGGATCTGGGGC	60°C
Ovine MAIT TCR Vα19 gene	boovUntr1f GCCTGGTTTGTGGTGTCCTG boovUntr2f CGCCTGGTTTGTGGTGTCCT boovUntr3f AGCGCCTGGTTTGTGGTGTC	boovAJ33r - boovAJ33r - boovAJ33r -	60°C 60°C 60°C
Ovine MAIT TCRα V+C regions	boovUntr3f -	CorovC3r AAGGTGGTCGGGTTTAACCT	60°C
Bovine MAIT TCRα V+C regions *	I GCTCTGCAGGAAAAGGCGTTAAG	Jr CCAAGCTGGTAGAGAAAAGCTTTGAAA	56°C
Confirmation of ovine MAIT TCRα CDR3	boovAV19f - boovAV19f - boovAV19f - boovUntr1f	ovTCRAJ1r ATCTGGGGCTCTGGGACC ovTCRAJ2r GACCAAGCTAATTATAAAGCCAG ovTCRAJ3r GCTCTGGGACCAAGCTA ovTCRAJ1r	56°C 56°C 56°C 56°C

*** These primers were designed by Dr Ildiko van Rhijn, University of Utrecht.**

Table 3-2: Primers used in PCR to amplify plasmid DNA extracted from clones derived from RACE-PCR product DNA, and optimal annealing temperature (Tm) for each reaction.

Target Sequence	Forward 5' sequence 3'	Reverse 5' sequence 3'	Optimal Tm
TCRα Constant Region	CLov5f ACCCCAACCCCACTGTGTA	BovCintr TGC TCA ACA TGG AGA TCC TG	58°C
Vα19 gene in combination with any Jα region	boovAV19f -	BovCintr -	60°C
MAIT Vα19-Jα33	boovAV19f -	boovAJ33r -	60°C
Jα33 gene in combination with any Vα region	RACE USP CTAATACGACTCACTATAGGGC	boovAJ33r -	64°C

3.2.4 PCR product cloning, transformation and selection for expansion

DNA from appropriate PCR products was either purified directly (see section 2.4.3.5) or extracted from bands in agarose gels (2.4.3.4). These DNA samples were cloned into pGEM-T Easy vector, and used to transform Promega JM109 high efficiency competent cells (see section 2.4.4). Transformed *E. coli* colonies were selected on the basis of antibiotic resistance and absence of beta galactosidase, and tested by colony PCR using either pGEM-T Easy primers, T7 and SP6, or primers specific to the insert.

3.2.5 Sequencing of plasmid DNA

Colonies of interest were grown overnight in LB medium and plasmid DNA extracted (see section 2.4.1.3). Sequencing of plasmid DNA from single clones was performed by DBS Genomics, Durham University, UK (see section 2.4.3.6). RACE-PCR product clones which were processed in 96-well plates were prepared and

sequenced in an alternate manner – colonies were denatured, inserts amplified by T7-SP6 PCR, and sequencing reactions performed within the 96-well plates using BigDye Terminator v3.1, using the protocol described in section 2.4.3.7. The specific sequencing primer used was based near the 5' end of the TCRα constant region: "BovCintr" – 5'-TGC TCA ACA TGG AGA TCC TG-3'. Sequencing was performed by the Sequencing Unit of the Department of Zoology, University of Oxford, UK.

3.2.6 RACE – primers and reaction conditions

The basis of RACE PCR, and general primers and cycling conditions, are described in section 2.4.2.3. 5' RACE PCR was used to amplify TCRα sequences from different fractions of T-cells purified by flow cytometry. The reverse gene-specific primer used for TCRα amplification was based in the middle of the TCRα constant region: BovC4r – 5'-TCC AAG AGC AAC GGG ATA GTG ACC TG-3'. RACE PCR products were assessed by agarose gel electrophoresis, purified, cloned into pGEM-T Easy vector and used to transform competent cells, as described for regular PCR products (sections 3.2.3 and 3.2.4).

3.2.7 Analysis and purification of cells by flow cytometry

Staining of cell suspensions by immunofluorescence for phenotyping and analysis by flow cytometry using a BD FACSCalibur Flow Cytometer (BD Biosciences, San Jose, California, USA) are covered in section 2.3.3. Primary monoclonal antibodies used for staining were MM1A for CD3, IL-A12 for CD4, IL-A105 for CD8, GB21A for γδ T-cells, as described in section 2.3.3.1, table 2-1. Cell subsets were purified by flow cytometry using a BD FACSAria Flow Cytometer (BD Biosciences, San Jose, California, USA) and FACSDiva software version 5.02. Purified cell fractions were re-analysed and shown to be 95-99% pure.

3.3 Results

3.3.1 Amplification of ovine and bovine orthologous MAIT TCRα chains

The human MAIT TCRVα chain sequence was used to search the NCBI online ovine gene databases, using the basic local alignment search tool, BLAST; no similar or orthologous sequences were found, indicating that the ovine MAIT TCRα chain had yet to be identified.

A partial bovine TCRα chain sequence orthologous to the human and murine MAIT TCRVα chains had previously been identified (Tilloy et al. 1999). These investigators had designed primers for amplifying the partial bovine TCRVα (including the CDR3 region) and these were replicated for this study as boovAV19f / boovAJ33r. Using an optimised PCR protocol at a range of annealing temperatures (60°C – 66°C), this primer pair was tested on both bovine and ovine PBMC cDNA (see figure 3-1). A strong, single product of expected size (120bp) was produced from bovine PBMC cDNA at all annealing temperatures, while a similar single product was produced from ovine cDNA at 60 and 62°C. A negative control using distilled water rather than cDNA showed no amplification, although some smearing on agarose gel electrophoresis suggested primer-dimer formation. To positively identify the PCR product as an ovine orthologue of MAIT TCRα, DNA was purified and cloned into the pGEM-T Easy vector. Screening of a number of white colonies, using boovAV19f / boovAJ33r (see figure 3-1) identified 4 clones (No's. 1, 2, 4 and 5) that yielded a product of the expected size suggesting the correct insert. Sequencing of the DNA inserts in these clones confirmed that the nucleotide sequences had close identity with previously described bovine MAIT TCRVα/CDR3 region sequence. Detailed analysis of the sequences is covered in section 3.3.2.

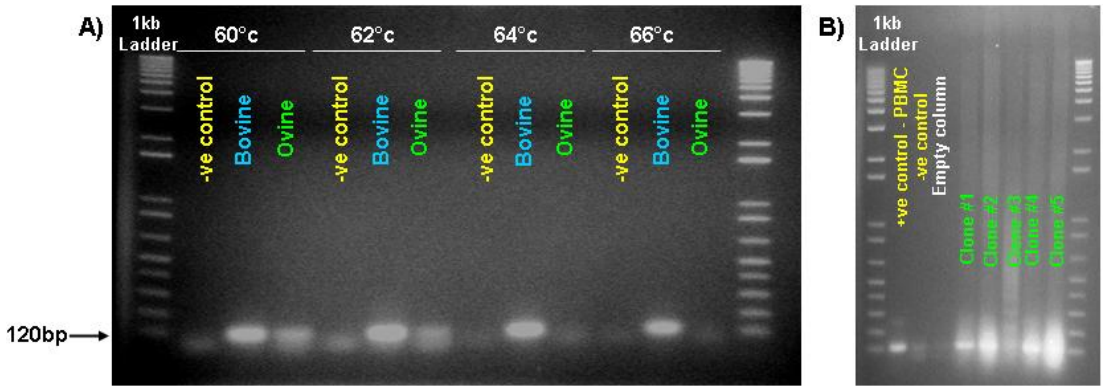


Figure 3-1: Agarose gels showing: A) Detection of 120bp PCR product, representing MAIT TCRα Vα19-Jα33 CDR3, amplified from bovine and ovine PBMC cDNA, using primers boovAV19f/boovAJ33r at a range of annealing temperatures. B) Positive amplification from most subclones of ovine Vα19-Jα33 PCR product, using the same primers and annealing temperature 60°C. Autoclaved distilled water was used as a negative control for both PCRs, and ovine PBMC was used as a positive control for the PCR in gel B.

To help identify the full-length variable gene and CDR3 sequences for the ovine MAIT TCRα chain, and since no matches had been found in the online ovine gene databases, the human MAIT Vα gene, Vα7.2, was used to search the NCBI bovine gene database. A total of 13 matching gene segments were identified. These were aligned to the human MAIT TCRα to identify a consensus sequence for the 5' untranslated region of the potential bovine MAIT TCRα chain. Based on this sequence, a series of forward primers were designed (boovUntr1f, boovUntr2f, boovUntr3f, boovUntr4f) for use with the reverse joining region primer, boovAJ33r. These primer pairs were tested on cDNA from both bovine and ovine PBMC, at a range of annealing temperatures (60°C, 63°C). Single, positive products of anticipated size (400bp) were produced from both bovine and ovine cDNA using primers boovUntr1f and boovUntr3f, the latter giving the best results when used at Tm:60°C (see figure 3-2). Products from ovine cDNA were purified and cloned into pGEM-T Easy vector, and clones were tested for insert by PCR. Sequencing of inserts from six clones with the correct anticipated insert size identified four clones with sequences displaying close identity to the previously defined partial bovine MAIT TCRVα sequence (further detailed in section 3.3.2).

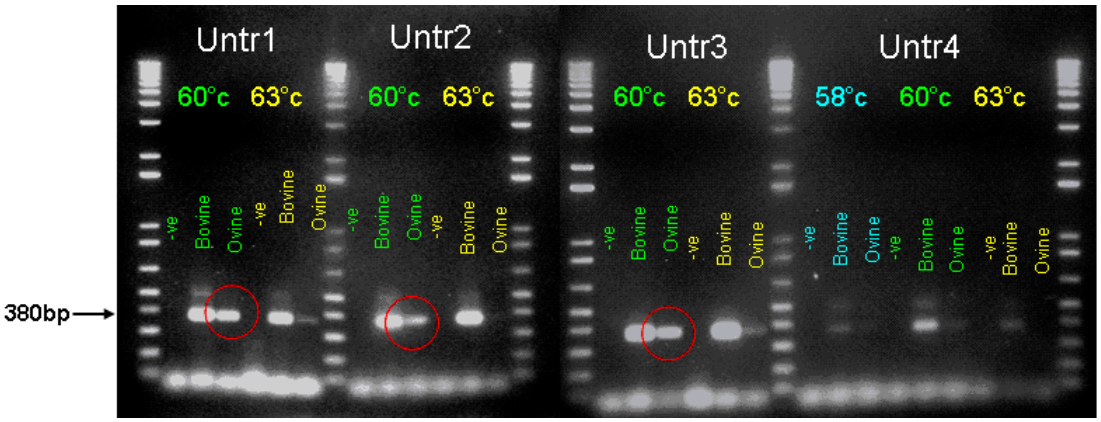


Figure 3-2: Agarose gel showing PCR amplification of 380bp product, representing the MAIT TCR Vα19 gene, using primers based in the Vα19 gene 5' untranslated region (boovUntr1f, boovUntr2f, boovUntr3f, boovUntr4f) and Jα33 (boovAJ33r), amplified from both bovine and ovine PBMC cDNA at a range of annealing temperatures. Positive results of anticipated size from ovine PBMC cDNA are circled in red. Autoclaved distilled water was used as a negative control for all PCRs.

Since these products were obtained using a reverse primer that extended into the CDR region, a further PCR assay using the same forward primer boovUntr3f with two newly designed ovine TCRα constant region reverse primers, CorovC1r and CorovC3r (located at the 3' end of the C gene), was conducted in an attempt to amplify the full-length variable and CDR3 sequences of the ovine MAIT TCRα chain. The primer pair boovUntr3f / CorovC3r produced a single positive band of the anticipated size (820bp) from ovine cDNA (see figure 3-3, A). Seven out of 12 cDNA clones examined from this PCR product were found to contain inserts of variable size, at, or near, 900bp, when screened by PCR using the same primers (see figure 3-3, B). Sequencing of the inserts from these clones identified two clones (clone #2 and #13) with TCRα sequences showing a high level of identity to the bovine MAIT TCRα chain, and identical to the partial Vα sequence produced by boovUntr3f / boovAJ33r (above). The other five clones yielded sequences

containing the same V gene segment (orthologous to human Vα7.2) but very different CDR3 sequences (further details in section 3.3.3).

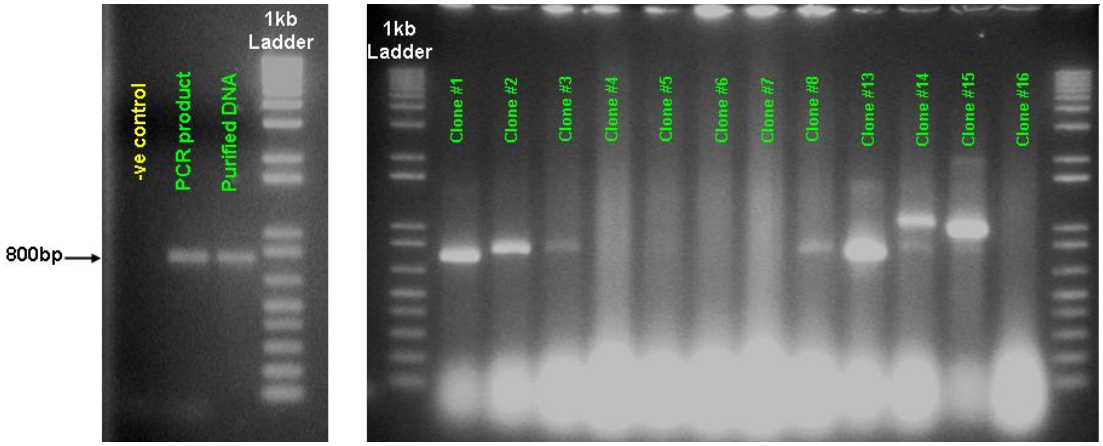


Figure 3-3: Agarose gels showing: A) Amplification, and purification, of a single product from PCR of ovine PBMC cDNA using primers based in the Vα19 gene 5' untranslated region (boovUntr3f), and the TCRα constant region (CorovC3r). Autoclaved distilled water was used as a PCR negative control. B) Sub-cloning of purified DNA revealed multiple different insert sizes between different colonies. Clones #2 and #13 returned MAIT Vα19-Jα33 sequence, while four other clones (#1, #3, #14, #15) returned sequence of Vα19 in combination with different Jα exons.

A similar approach was used to obtain the complete sequence of the bovine MAIT Vα and CDR3 coding region. PCR primers "I" and "Jr", based in the 5' untranslated region and constant region, respectively, had been concurrently designed by Dr Ildiko van Rhijn, University of Utrecht, and were tested on cDNA from bovine PBMC at a range of annealing temperatures (56 – 62°C). Strongest amplification of product of the anticipated size (650bp) was at 56°C, although a second smaller band was also amplified at 56°C and 59°C. The 650bp band was excised from the agarose gel, and DNA extracted and cloned into the pGEM-T Easy vector. Sequencing of inserts from four clones with the anticipated insert size revealed three almost identical sequences; one of these displayed 100% identity to the previously defined

partial bovine MAIT TCRVα sequence, Vα19-Jα33, confirming positive amplification of the bovine MAIT Vα gene, while sequence from two other clones each showed only a single nucleotide difference in the CDR3 region in comparison with the first clone. Sequence from the fourth clone contained Vα19 in combination with a different Jα gene (Jα20) and had dramatically altered CDR3 sequence (see section 3.3.4). Detailed discussion of the sequences is presented in section 3.3.2.

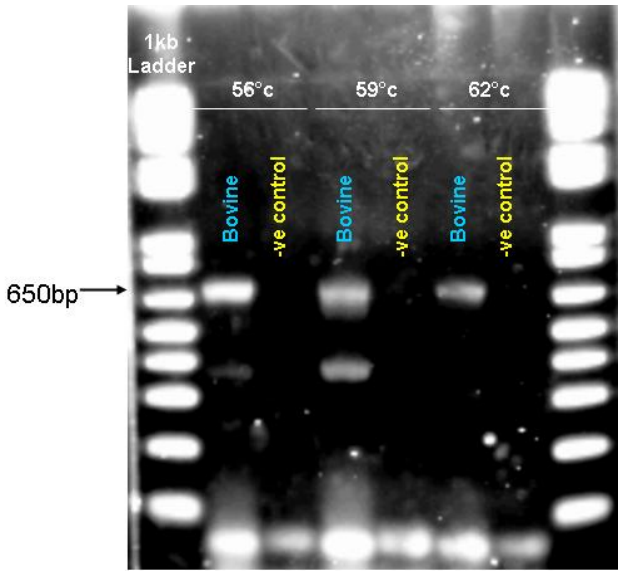


Figure 3-4: Agarose gel showing amplification of 650bp PCR product representing bovine MAIT TCRα from bovine PBMC cDNA using primers I/Jr, based in the 5' untranslated region of Vα19, and the TCRα constant region, respectively, at a range of annealing temperatures. Autoclaved distilled water was used as a PCR negative control. DNA extracted from the 650bp size bands was subcloned, and 3 out of 4 clones were sequenced as Vα19-Jα33; the fourth clone was Vα19-Jα20.

3.3.2 Variation of CDR3 region coding sequence and high level of cross-species similarity

Sequence data obtained with different primer combinations on ovine PBMC cDNA all contained identical or near identical TCRVα gene segment sequences, but showed differences within the CDR3 region. Sequences obtained using the reverse primer boovAJ33r, which extended into the CDR3, differed by 2 nucleotides from those amplified using a reverse primer in the constant region (CorovC3r), resulting in two amino acid substitutions in the predicted CDR3 amino acid sequences: M D G N Y Q W I compared to M D G N Y R L I, respectively. To test whether this discrepancy was introduced by primer boovAJ33r, which was sited over the 3' end of the CDR3 region, 3 different reverse primers in the 3' end of bovine Jα33 (ovTCRAJ1r, ovTCRAJ2r, and ovTCRAJ3r, all 3' of the region showing sequence discrepancy) were paired with different Vα forward primers (boovAV19f, boovUntr1f, and boovUntr3f) to amplify specifically this segment of the MAIT α chain from ovine cDNA (see figure 3-5 for primer location summary).

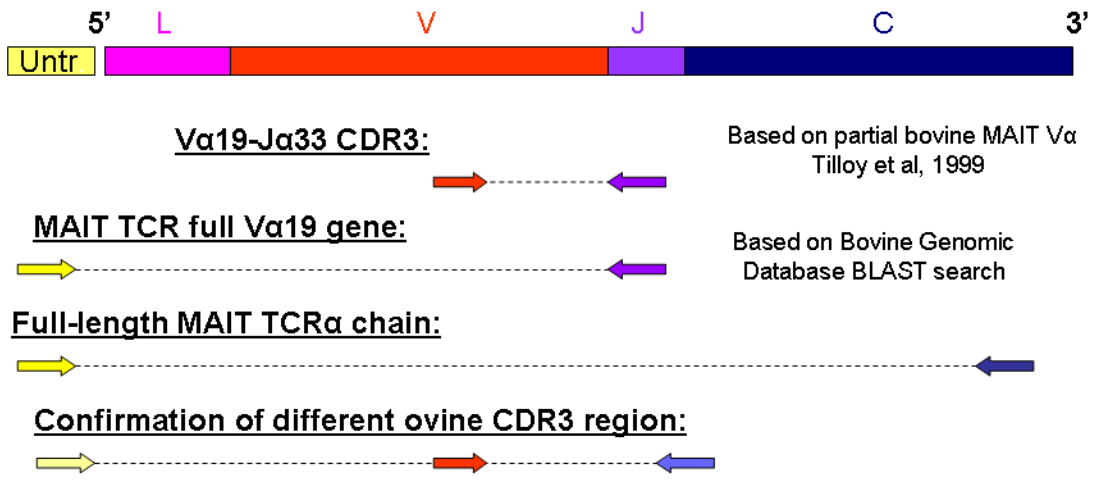


Figure 3-5: Summary of primer locations used to identify bovine and ovine MAIT TCRVα gene sequence. Untr: 5' untranslated sequence; L: Leader sequence; V: Vα19 gene; J: Jα33 gene; C: Constant gene.

Products of expected size were amplified by boovAV19f with each reverse primer, as well as from boovUntr1f and ovTCRAJ1r. Sequencing of cDNA clones from each of these four products yielded the same nucleotide sequence, which was identical to that obtained above using a reverse PCR primer located in the constant region. These results confirmed that the bovine and ovine MAIT CDR3 region sequences differ at 2 nucleotide positions, resulting in two amino acid substitutions near the 3'end of the CDR3 region (see figure 3-6).

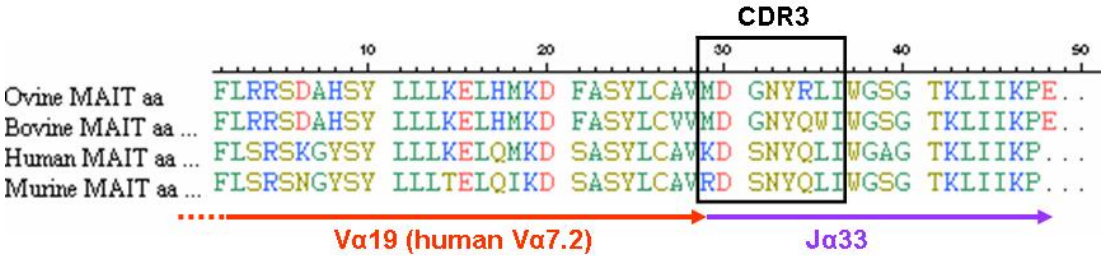


Figure 3-6: Cross-species comparison of the CDR3 of the potential polypeptides encoded by ovine, bovine, human and murine MAIT TCRVα sequences. The CDR3 region is highlighted and demonstrates cross-species conservation, but also differences between the 3' amino acids of bovine and ovine CDR3 sequences.

Sequence analysis of bovine and ovine MAIT TCRα chain sequences revealed a further intra-species discrepancy at the 5' end of the CDR3 region, near the site of the junction between V and J gene segments. In both species, the same single base pair substitution resulted in a coding change of the first CDR3 codon from methionine to isoleucine. This change was seen in three out of ten bovine sequences and three from fifteen ovine sequences. When these sequences were compared to the corresponding bovine genomic V (Vα19) and J (Jα33) gene segment sequences (sourced online from Ensembl and confirmed with IMGT databases), the two bovine cDNA sequences can be accounted for by differential nucleotide deletions at the VJ junction, with in one case deletions at the 3' end of the V gene generating a ATG

codon primarily from the J segment (encoding methionine), and in the other case deletions at the 5' end of the J segment resulting in a ATA codon from the V gene (encoding isoleucine) (see figure 3-7). Due to high nucleotide identity between cattle and sheep TCRα gene sequences, it is proposed that a similar mechanism accounts for the same substitution observed in the ovine translational sequence. Due to its closer identity to human and murine MAIT TCRα chains, the bovine and ovine MAIT TCRα chain sequences containing methionine at this position were used for subsequent cross-species sequence comparisons.

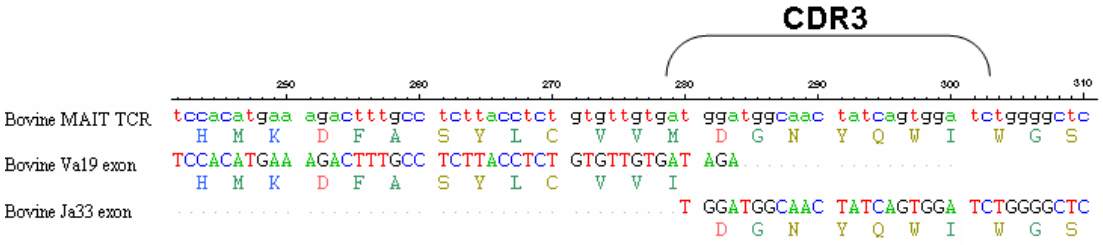


Figure 3-7: Bovine germline sequence for Vα19 and Jα33 gene segments overlap at the site of the MAIT TCRα CDR3. Final MAIT TCRα sequence must be a result of nucleotide deletion. Although consensus bovine CDR3 sequence starts with codon ATG encoding methionine due to deletions at the 3' end of Vα19 germline sequence, some sequences have codon ATA due to deletions at the 5' end of Jα33 germline sequence, which encodes for isoleucine. Both of these sequences are represented in cDNA from PBMC.

MAIT TCRα chain sequences from different species showed a high level of conservation at both the nucleotide and amino acid level (see table 3-3). The levels of nucleotide similarity between the ovine, bovine, murine and human MAIT TCRα germline Vα gene elements was >74%, consistent with them belonging to orthologous V gene subfamilies, and was >83% for the ovine, bovine and human Vα sequences. The Jα33 germline sequence, which encodes the majority of the CDR3 region, was highly conserved between the 4 species, at over 91% nucleotide identity (see Appendix C for full cross-species nucleotide sequence comparison). These

similarities were reflected by a high level of amino acid identity between the ovine, bovine, murine and human MAIT TCRα (V-region) transcripts, which was >70% for all species, and was 93% for the two ruminant species. The CDR1 and CDR2 regions within the Vα sequence were markedly conserved: CDR1 has 100% amino acid identity between all four species, and the CDR2 of ovine, bovine and human sequences are identical, while the murine CDR2 contains a single amino acid substitution within the first codon. The CDR3 regions are also highly conserved between species, with identical length of 8 amino acids, and conservation of the second, fourth, fifth and eighth amino acids between all four species. However, a small number of conservative substitutions are seen in the first, third, sixth and seventh amino acids between species (see figure 3-8).

Table 3-3: Cross-species conservation of the MAIT Vα19 (Vα7.2 in humans) and Jα33 gene nucleotide sequences, and of the whole V- J- amino acid sequences, including CDR regions.

	Germline Vα19	Germline Jα33	Vα19-Jα33 protein identity			
Species	Nucleotide	Nucleotide	Amino Acid	CDR1	CDR2	CDR3
Ovine:Bovine	273/283 bp (96%)	54/57 bp (95%)	103/111 aa (93%)	6/6 aa	5/5 aa	6/8 aa
Ovine:Human	238/283 bp (84%)	53/57 bp (93%)	82/111 aa (74%)	6/6 aa	5/5 aa	5/8 aa
Ovine:Murine	213/283 bp (75%)	55/57 bp (96%)	79/111 aa (71%)	6/6 aa	4/5 aa	5/8 aa
Bovine:Human	234/283 bp (83%)	52/57 bp (91%)	82/111 aa (74%)	6/6 aa	5/5 aa	5/8 aa
Bovine:Murine	209/283 bp (74%)	54/57 bp (95%)	79/111 aa (71%)	6/6 aa	4/5 aa	5/8 aa
Human:Murine	210/283 bp (74%)	55/57 bp (96%)	78/111 aa (70%)	6/6 aa	4/5 aa	7/8 aa

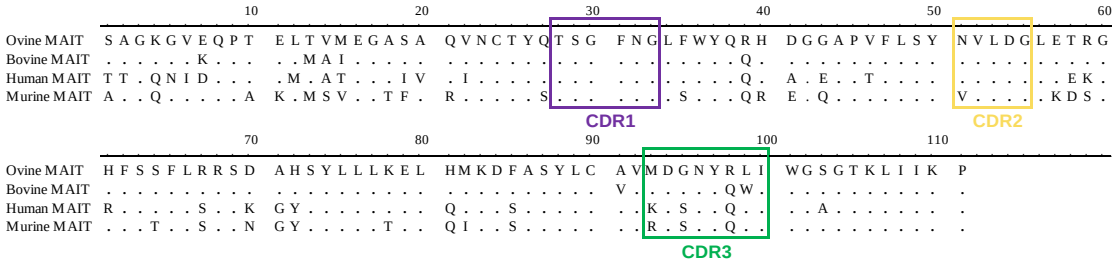


Figure 3-8: Cross-species comparison of ovine, bovine, human, and murine MAIT TCRα V-region coding sequences, showing cross-species conservation of the CDR1 (purple box), CDR2 (yellow box) and CDR3 (green box) regions.

3.3.3 No evidence of enrichment for MAIT cells within the CD4$^-$ CD8$^-$ T cell subset

Previous studies in humans and mice have indicated that MAIT cells are highly enriched in the small subset of CD4$^-$ CD8$^-$ T cells. An experiment was undertaken to investigate whether this also applies to cattle. Firstly, the number of CD4$^-$ CD8$^-$ αβ T cells in bovine PBMC was assessed by two-colour immunofluorescence staining and flow cytometry, using an antibody specific for CD3 against a mixture of 3 antibodies specific for CD4, CD8 and the γδ TCR (see figure 3-9). The number of CD4$^-$ CD8$^-$ αβ T-cells in PBMC from 6 cattle ranged between 1.9% and 3.8% (see table 3-4). One of these animals (332) was used for further analysis of MAIT TCRα expression.

Table 3-4: Flow cytometry analysis of PBMC from six cows identified the $CD4^-/CD8^-$ αβ T-cell population as representing 1.9-3.8% of PBMC. Animal 332 had relatively increased $CD4^-/CD8^-$ αβ T-cell population and high cell yield and was used for further analyses.

Animal	Cells/ml obtained from PBMC	$CD4^-/CD8^-/\gamma\delta^-$ T-cell Population %
002	1.2×10^7	3.4%
332	2.3×10^7	3.6%
475	1.5×10^7	1.9%
481	1.6×10^7	3.0%
851	1.9×10^7	3.8%
860	2.7×10^7	2.3%

PBMC from this animal were stained as above by immunofluorescence, and $CD3^+$ $CD4^+/CD8^{+/}\gamma\delta^+$ and $CD3^+$ $CD4^-$ $CD8^-$ $\gamma\delta^-$ populations purified by cell-sorting, in order to analyse the expressed TCRα chains. The staining profile and gates set for sorting are shown in figure 3-9. The resultant populations were 99% and 95% pure, respectively. Staining for γδ TCR was included to ensure that the $CD4^-$ $CD8^-$ αβ T-cell population was not contaminated with γδ T cells. RNA was extracted from the two sorted populations and from a sample of the unfractionated PBMC.

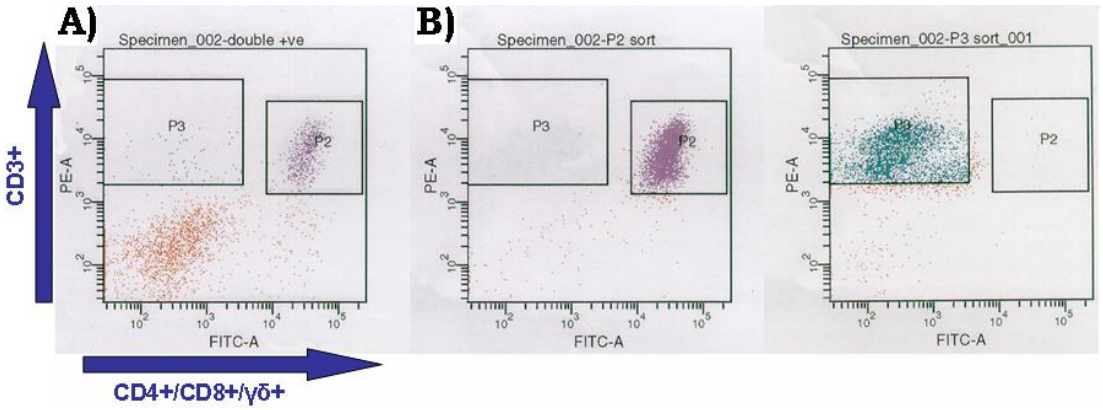


Figure 3-9: A) Flow cytometry purification of bovine PBMC, where $CD3^+$ cells stain PE^+, while $CD4^+$, $CD8^+$, and $\gamma\delta^+$ cells stain $FITC^+$. Cells were sorted into gated-populations P2 ($CD3^+$, $CD4^+/CD8^+/\gamma\delta^+$) and P3 ($CD3^+$, $CD4^-$, $CD8^-$, $\gamma\delta^-$). B) Testing for purity of each population (99% for sample P2, 95% for sample P3).

These three samples were subjected to 5' RACE PCR to amplify a representation of the TCRα transcripts within each sample. RACE-ready cDNA was synthesised from RNA through reverse transcription using Moloney murine leukaemia virus reverse transcriptase. The dC-tailing action of this enzyme allowed annealing of an additional oligonucleotide to the 5' end of the original RNA sequence, which served as an extended template for further cDNA synthesis. RACE-ready cDNA was amplified by PCR with a universal primer mix specific for the oligonucleotide sequence added to the RNA template, and a reverse TCRα-chain-specific primer (BovC4r) based in the bovine constant region (see figure 3-10). The RACE PCR products gave a broad band of between 500 and 1000 bp when run in agarose gels (see figure 3-11).

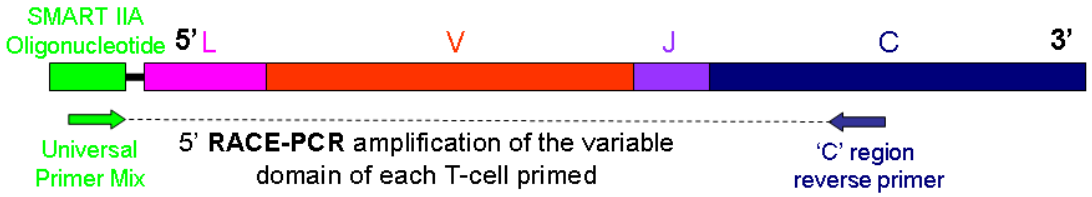


Figure 3-10: Summary of application of Rapid Amplification of cDNA Ends (RACE)-PCR for amplification of a range of different T-cell variable regions from cDNA. L: leader sequence; V: any Variable gene; J: any Joining gene; C: Constant gene.

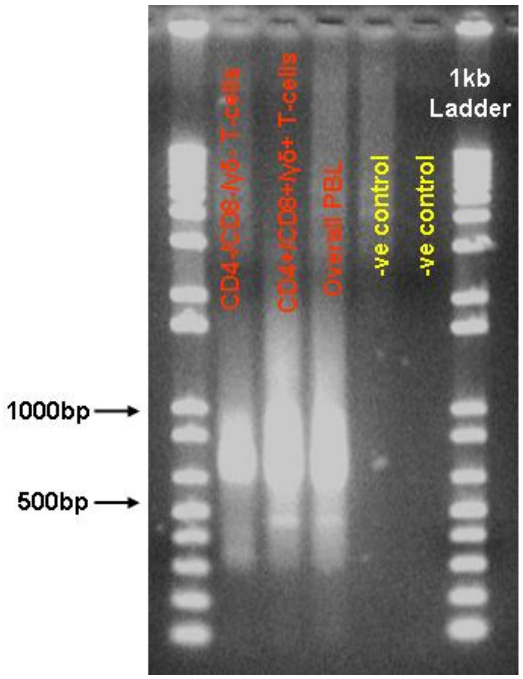


Figure 3-11: Agarose gel showing amplification of a broad band of PCR product (500-1000bp), by 5' Rapid Amplification of cDNA Ends (RACE)-PCR using a gene-specific reverse primer based in the TCRα Constant region, from each of three samples, $CD4^-/CD8^-/\gamma\delta^-$ T-cells, $CD4^+/CD8^+/\gamma\delta^+$ T-cells, and overall PBL. Two negative controls, lacking either the forward (RACE USP) or reverse (BovC4r) primers, respectively, were included.

The DNA products of the three RACE reactions were purified and cloned into pGEM-T Easy vector. One hundred and eighty colonies with inserts (white colonies)

were amplified and added to the wells of two 96-well plates (90/plate), with 6 wells in each plate containing nuclease-free water serving as negative controls. Plasmids were denatured and insert-DNA amplified in the wells by vector-specific primers, before being processed for sequencing using another TCRα-specific primer (BovCintr) based in the bovine constant region. The quality of nucleotide sequences obtained from these multiple-well sequencing reactions was highly variable; usable sequences were obtained from 89, 115, and 124 clones derived from PBMC, $CD4^-$ $CD8^-$ $\gamma\delta^-$, and $CD4^+$+$CD8^+$+$\gamma\delta^+$ samples, respectively. Negative control wells did not produce sequence results. Sequences were BLAST-searched through an internal database of human TCRVα sequences established by Dr Tim Connelley, University of Edinburgh. Since some of the sequence data produced was not of good quality and only contained short areas of high quality sequence in the expected location for Vα sequence, positive BLAST matches of >50bp length and >90% nucleotide similarity were taken as an estimation of the equivalent human TCRVα subfamily of DNA from each colony. The high percentage of nucleotide similarity was set to compensate for the inclusion of short areas of matching sequence. The numbers of sequences matching different human TCRVα subfamilies obtained from each cell fraction are displayed in figure 3-12. Overall, thirty different human TCRVα subfamilies were represented. Some subfamilies appeared to be represented more frequently and there was some variation in subfamily representation between the samples but the overall numbers of samples examined were not sufficient to make statistically meaningful comparisons. Nevertheless, the TCRVα7 subfamily, which in humans contains two gene members, one of which is expressed by MAIT cells, was represented 4 times in 89 cDNA clones from unfractionated PBMC, but not in clones from the $CD4^+$+$CD8^+$+$\gamma\delta^+$ or $CD4^-$ $CD8^-$ $\gamma\delta^-$ populations. Further sequence analysis confirmed that these four sequences were an exact match to the previously identified bovine MAIT Vα19, and were all used in combination with Jα33. These findings did not reveal any obvious enrichment for MAIT TCRα chain sequences within the $CD4^-$ $CD8^-$ T cell population.

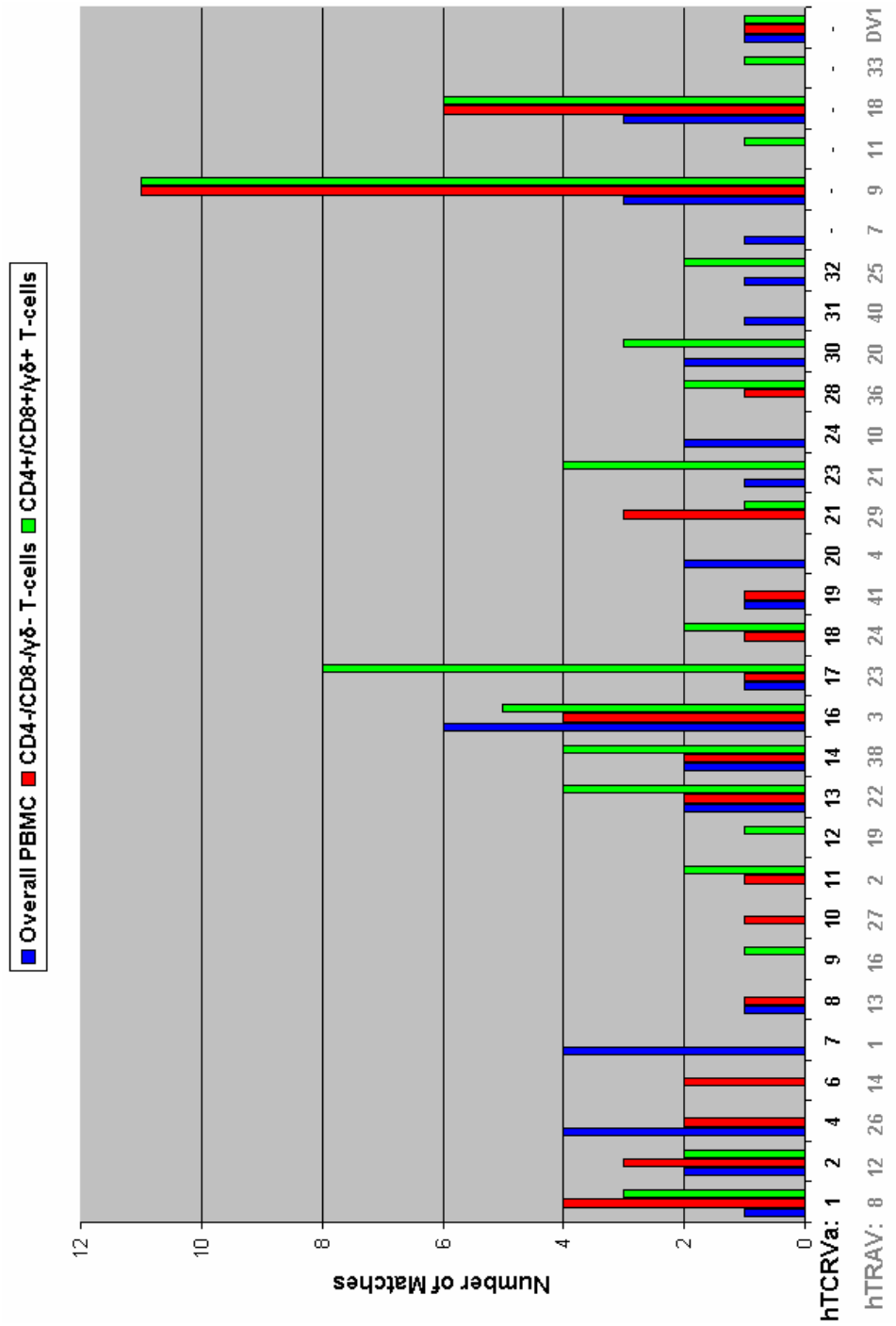


Figure 3-12: Summary of range of TCRVα subfamilies represented within each of three samples, bovine PBMC (blue), $CD4^-$ $CD8^-$ $\gamma\delta^-$ T-cells (red) and $CD4^+$+$CD8^+$+$\gamma\delta^+$ T-cells (green). 5' RACE PCR was used to amplify the different TCRVα genes within each sample, and product DNA was subcloned. Clone inserts were sequenced and identified by homology to subfamilies of human TCRVα genes. Four matches to the human MAIT Vα7.2 were found in whole PBMC, but none in the other phenotypic subsets. 5 sequences could not be matched to human WHO-IUIS sequence nomenclature and so were identified using the alternative IMGT nomenclature.

3.3.4 Non-exclusive usage of MAIT TCRα V and J gene segments within the TCRα repertoire

The frequency of occurrence of the bovine MAIT variable TCRα gene segments, Vα19 and Jα33, amongst the TCRα cDNA clones generated by RACE-PCR was investigated. Plasmid DNA from each of the previous sets of cDNA clones, unfractionated PBMC, $CD4^+$+$CD8^+$+$\gamma\delta^+$ and $CD4^-$ $CD8^-$ $\gamma\delta^-$ T-cells, was tested by PCR with three primer pairs: (i) CLov5f / BovCintr - to amplify TCRα constant region and hence confirm the presence of an α chain; (ii) boovAV19f / BovCintr - to amplify sequences with Vα19 in combination with any Jα gene; (iii) boovAV19f / boovAJ33r - to amplify the specific MAIT TCRα chain containing Vα19 and Jα33. Each PCR was applied to 180 clones from PBMC (plus 12 negative control wells), and 90 clones each from the $CD4^+$+$CD8^+$+$\gamma\delta^+$ and $CD4^-$ $CD8^-$ $\gamma\delta^-$ populations (each with 6 negative control wells). Examples of results obtained with these PCRs in the cDNA clones from PBMC are shown in figure 3-13 and a summary of results from each PCR reaction is presented in figure 3-14.

Most clones were positive for TCRα constant region: 168/180 for PBMC, 89/90 for $CD4^-$ $CD8^-$ $\gamma\delta^-$ population and 86/90 for the $CD4^+$+$CD8^+$+$\gamma\delta^+$ T-cells. In clones from PBMC, a strong positive product for Vα19 was detected in 4 clones (from a total of 168 clones positive for TCRα) and all of these clones also gave a positive result with the PCR specific for the MAIT sequence, Vα19-Jα33. Sequencing of these 4 clones confirmed that they all contained the bovine MAIT Vα gene. A number of the clones from the $CD4^-$ $CD8^-$ $\gamma\delta^-$ (8 clones) and $CD4^+$+$CD8^+$+$\gamma\delta^+$ (5 clones) T-cell populations gave positive products for Vα19, but none of these clones gave positive results for MAIT TCRα. Sequencing of two of these clones ($CD4^+$+$CD8^+$+$\gamma\delta^+$ B2 and G10) revealed that they did not contain Vα19, but contained Vα22-Jα17 and Vα18-Jα17 respectively, demonstrating that the boovAV19f primer is not completely specific for Vα19. Sequencing of a number of weak or incorrectly sized bands obtained with the boovAV19f – boovAJ33r PCR showed that none of them represented Vα19-Jα33.

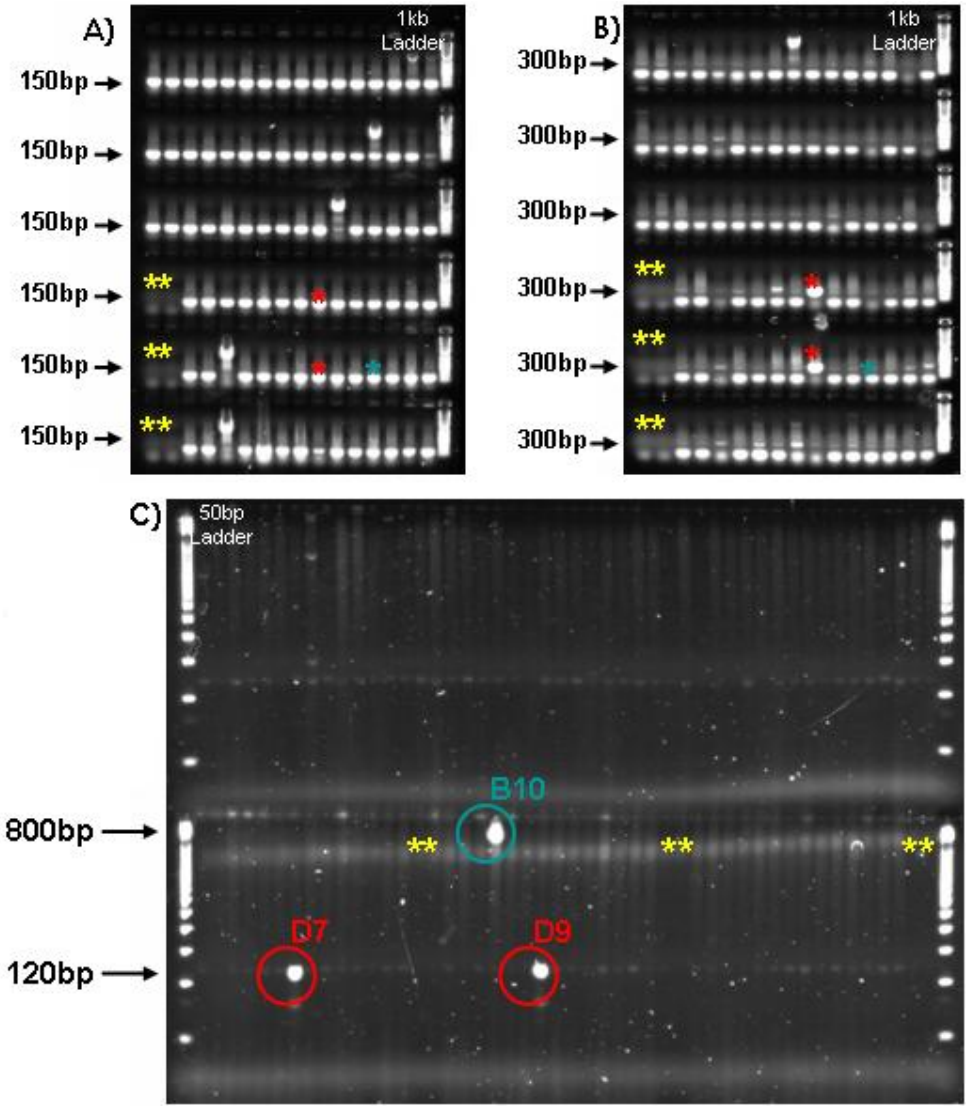


Figure 3-13: Agarose gels showing amplification of (A) TCRα constant region (150bp), (B) Vα19 gene (300bp, N.B. all sample columns show some formation of primer dimer), (C) MAIT TCRVα Vα19-Jα33 (120bp), in PCR of plasmid DNA from clones of different TCRVα chains amplified in 5' RACE PCR from overall PBMC (plate#1 of 2 plates). Primers used were: A) CLov5f-BovCintr, B) boovAV19f-BovCintr, C) boovAV19f-boovAJ33r. Wells D7 and D9 are positive for MAIT TCRVα (120bp), and their respective positive bands for constant region and Vα19 are marked by red asterisks; sequencing confirmed that both clones contained Vα19-Jα33. Well B10 amplified a large unexpected product (800bp) with primers boovAV19f – boovAJ33r, and was positive for TCR constant region, but did not amplify product representing the Vα19 gene (blue asterisks); sequencing revealed this clone to contain Vα18 and Jα33. Six negative controls consisting of autoclaved distilled water were included for each reaction (marked by yellow asterisks).

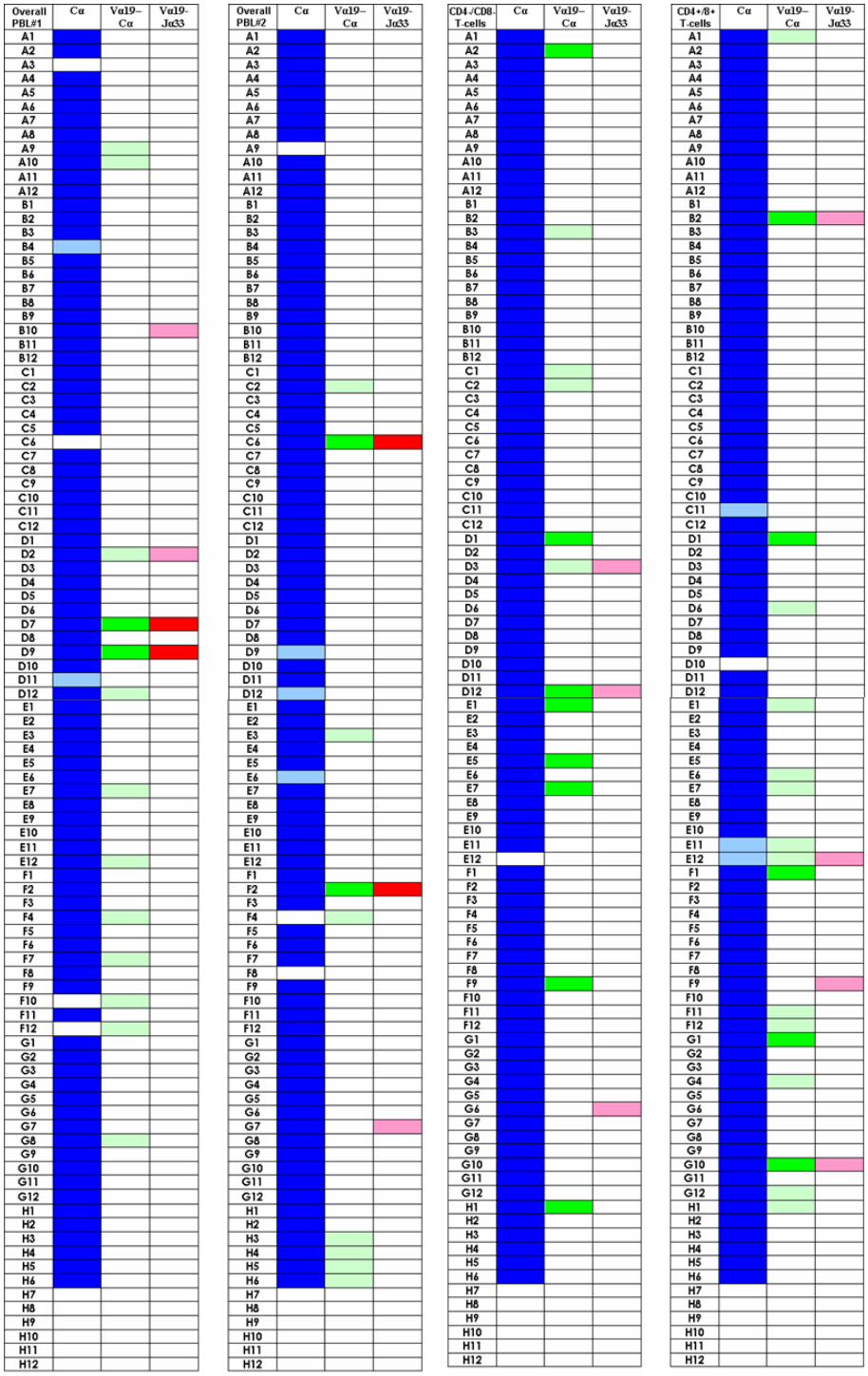


Figure 3-14: Positive (full colour), weak/incorrect (pale colour), or negative (blank) amplification of TCRα constant region (blue), Vα19 but unspecific Jα (green), and Vα19-Jα33 (red) product from DNA of RACE-amplified TCRα sequences. From PBMC, 4 clones contained Vα19-Jα33 sequence from 168 clones positive for TCRα constant region (columns 1 & 2); no Vα19-Jα33 sequences were detected in $CD4^-$ $CD8^-$ $\gamma\delta^-$ (column 3) or $CD4^+$+$CD8^+$+$\gamma\delta^+$ T-cell subset clones (column 4), from 89 and 86 clones positive for TCRα constant region, respectively.

Since the Jα33 gene segment contributes the majority of sequence to the MAIT TCRα CDR3 region, the sequence data obtained in 3.3.3 above was blasted with the bovine Jα33 sequence to identify clones expressing this J gene segment and to examine the relationship of Jα33 and Vα19 expression. The Jα33 gene was identified in 6 out of 328 clones for which sequence data were available; 4 of these clones also incorporated Vα19 gene segment to form the MAIT TCRα chain (as previously identified – section 3.3.3). The other two Jα33^{+} clones, also from unfractionated PBMC, contained two different TCRVα genes, which matched to human Vα30 and Vα18 respectively (see table 3-5). The CDR3 regions of these two clones were both distinct from MAIT CDR3, and each other, due to changes in length and N-region coded amino acids (see figure 3-16).

To further investigate this question, the cDNA that had been obtained from PBMC by RACE PCR was subjected to PCR amplification using the RACE reaction universal short forward primer (USP) in combination with the Jα33-specific reverse primer (boovAJ33r) at three different annealing temperatures: 56°C, 59°C, 62°C. Each PCR product appeared as a broad band of the expected size (450bp) and an additional 600bp band on agarose gel electrophoresis. DNA from these were extracted and cloned separately into pGEM-T Easy vector. The resultant clones demonstrated different insert sizes within the expected range (see figure 3-15). Seven clones from the 450bp band selected for sequencing all contained Jα33, but in combination with a number of different Vα genes – three of them contained the MAIT sequence, Vα19-Jα33; the other four clones each contained different V gene segments - three matching human TCRVα 2, 9, and 22. The fourth sequence was found to bear high similarity to the IMGT TRAV18 subgroup; this could not be annotated using the WHO-IUIS nomenclature applied to this thesis, and is described using IMGT nomenclature – see section 2.5.3. The sequence of a single clone of the 600bp PCR product contained a V region of increased length (550bp compared to approx. 400bp for most other Vα regions identified) which showed close similarity to human Vα32, in combination with Jα33 (see table 3-5).

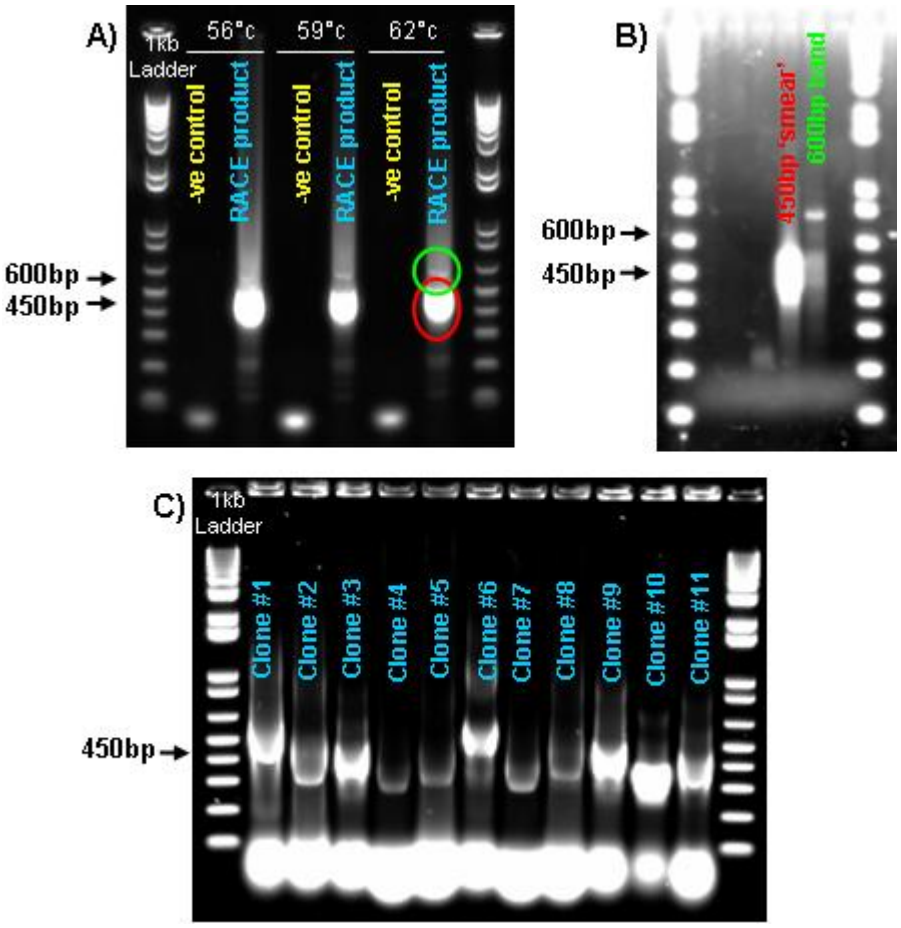


Figure 3-15: Agarose gels showing A) PCR amplification of RACE-PCR product TCRα sequences containing the Jα33 gene, using RACE USP primer and boovAJ33r at a range of annealing temperatures; these produced a broad band of 450bp product (red circle) and a discrete 600bp band (green circle); B) DNA extracted from the 450bp band and 600bp band in gel A; C) PCR screening of clones containing DNA from the 450bp band, using RACE USP and boovAJ33r primers, reveals a range of different insert sizes, representing different TCRα chains formed by use of various TCRVα genes in combination with Jα33.

The CDR3 regions of identified MAIT sequences were compared to those of clones containing the same J gene segment but different Vα gene segments. These showed at least one, and up to four, codon differences, primarily at or near the 'N region' at the 5' end of the CDR3. CDR3 length varied between 8 and 11 amino acids.

Examples of formed CDR3 sequences are shown in figure 3-16. The CDR1 and CDR2 sequences of each associated V gene were very different in comparison to those in the MAIT V, suggesting that the TCR is unlikely to be performing the same function despite similarities in CDR3.

Table 3-5: Vα genes used in combination with Jα33, in clones of RACE-product from bovine PBL ("332…"), or amplified by primers USP-boovAJ33r in PCR from RACE-product DNA ("450/600…").

	Match to:	
Sequence	**Human TCRVα**	**Human TCRJα**
332/1 B10	Vα 18	Jα 33
332/1 D7	Vα 7.2	Jα 33
332/1 D9	Vα 7.2	Jα 33
332/2 C6	Vα 7.2	Jα 33
332/2 F2	Vα 7.2	Jα 33
332/2 G7	Vα 30	Jα 33
450 clone#1	TRAV18(*)	Jα 33
450 clone#2	Vα 22	Jα 33
450 clone#3	Vα 7.2	Jα 33
450 clone#6	Vα 2	Jα 33
450 clone#9	Vα 7.2	Jα 33
450 clone#10	Vα 9	Jα 33
450 clone#11	Vα 7.2	Jα 33
600 clone#5	Vα 32	Jα 33

Matches to bovine MAIT TCRα Vα19-Jα33 are shown in blue - these composed 7 of 14 sequences. (*) Vα sequence in 450clone#1 did not match any annotated WHO-IUIS human TCRVα sequence, but was identified as a match to IMGT hTRAV18.

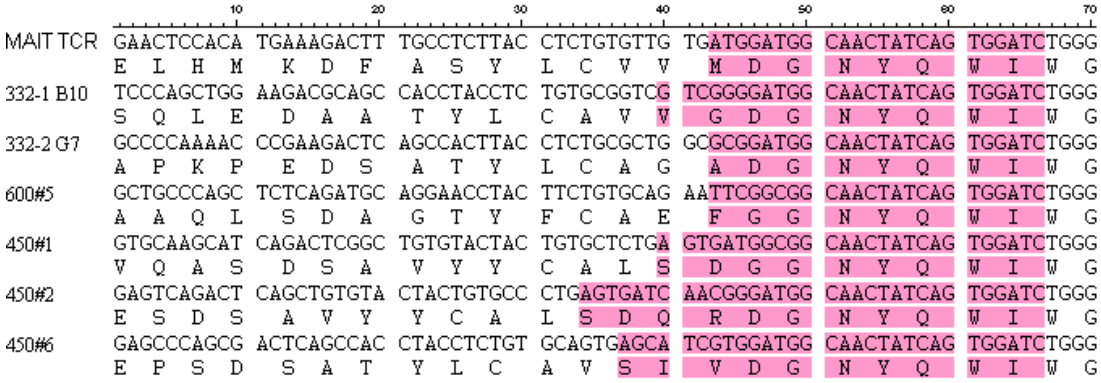


Figure 3-16: Nucleotide and amino acid sequences of clones of bovine TCRVα regions amplified by 5' RACE PCR, which contain Jα33 in combination with different Vα genes, showing the CDR3 sequences (highlighted pink) in comparison to the bovine MAIT CDR3 ("MAIT TCR"). CDR3 lengths vary between 8 and 11 amino acids and 5' clone CDR3 sequences are in some cases dramatically altered in comparison to the MAIT TCRα.

These results showed that use of the Jα33 gene is not specific to Vα19 and detailed changes in the CDR3 where Jα33 was used in combination with different Vα genes. Similarly, previous experiments in this chapter had identified that use of the MAIT Vα19 gene was not specific to Jα33, and identified Vα19 in combination with different Jα genes in both cattle and sheep. Sequences from these experiments were compared to the same human TCRVα database to identify genes used (see table 3-6), and the CDR3 of these sequences were also studied in comparison to the MAIT CDR3 (figure 3-17). As expected, since Jα33 is responsible for the majority of the MAIT CDR3, these sequences contained substantial diversity in CDR3 sequence and length (8 – 11 amino acids). Bovine sequence representing Vα19-Jα20 contained a markedly different CDR3 sequence with substantial nucleotide substitution within the 3' end of Vα19. Two ovine sequences, composed of Vα19-Jα50 and Vα19-Jα24, encoded stop codons with the CDR3 region, and would not encode functional TCRα chain protein.

Table 3-6: Jα genes used in combination with Vα19 in subclones of bovine PBMC amplified by primers I – Jr (IJr…) and ovine PBMC amplified by primers Untr3f – CorovC3r (U3C3…). Ovine/bovine MAIT TCRα chain orthologues are highlighted blue.

		Match to:	
Species	**Sequence**	**Human TCRVα**	**Human TCRJα**
Bovine	**IJr clone#1**	**Vα 7.2**	**Jα 33**
Bovine	**IJr clone#4**	**Vα 7.2**	**Jα 20**
Bovine	**IJr clone#7**	**Vα 7.2**	**Jα 33**
Bovine	**IJr clone#8**	**Vα 7.2**	**Jα 33**
Ovine	**U3C3 clone#1**	**Vα 7.2**	**Jα 23**
Ovine	**U3C3 clone#2**	**Vα 7.2**	**Jα 33**
Ovine	**U3C3 clone#3**	**Vα 7.2**	**Jα 35**
Ovine	**U3C3 clone#13**	**Vα 7.2**	**Jα 33**
Ovine	**U3C3 clone#14**	**Vα 7.2**	**Jα 50**
Ovine	**U3C3 clone#15**	**Vα 7.2**	**Jα 24**

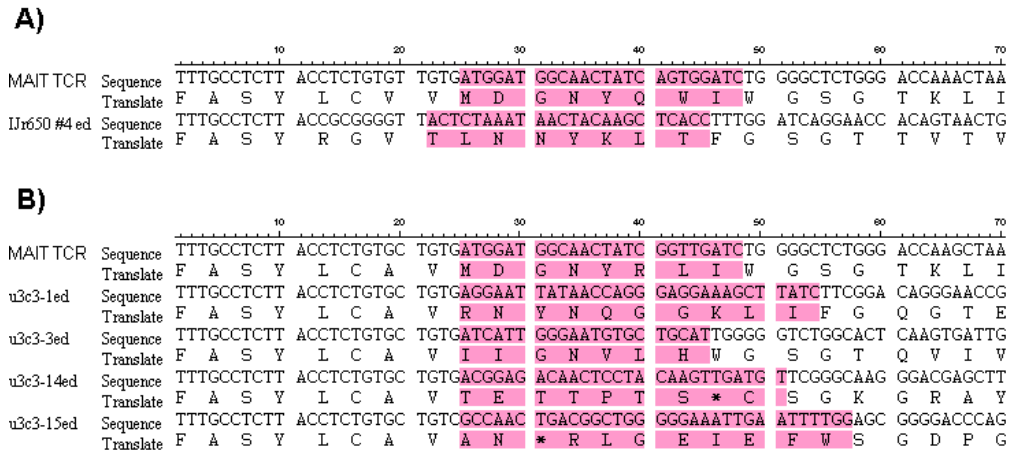


Figure 3-17: Nucleotide and amino acid sequences of clones of (A) bovine and (B) ovine TCRVα regions which contain Vα19 in combination with different Jα genes, showing the CDR3 sequences (highlighted pink) in comparison to the MAIT CDR3 for each species ("MAIT TCR"). These sequences were identified from clones of product of PCR using primers I/Jr (A) or boovUntr3f/CorovC3r (B) on cDNA from bovine and ovine PBMC, respectively. CDR3 lengths vary between 8 and 11 amino acids. Bovine clone sequence IJr650#4ed shows nucleotide substitution within the 3' Vα19 sequence. Ovine clone sequences u3c3-14ed and u3c3-15ed both contain a stop codon within the CDR3.

3.3.5 Bovine Vα19 is located at the start of the TCRα locus on chromosome 10

Searches of the bovine genome database using the MAIT TCRα Vα19 and Jα33 sequences confirmed the genomic location of the bovine TCRα locus, including the MAIT TCRα genes, to be on chromosome 10, as previously described (Reinink and Van Rhijn 2009). Bovine Vα19 mapped to chromosome 10 at 24.94Mb. The assembly of this part of the locus is not complete; however, bovine putative olfactory receptor genes (accession numbers: IPI00906644, IPI00911684, and IPI00907989) were located on either side of the BLAST location of Vα19 at 24.95Mb, 24.91Mb, and 24.90Mb respectively. A large number of "Ig-like" genes were identified downstream from this site, the first of which mapped to 24.86Mb. These had been previously identified as TCRVα genes by Reinink and Van Rhijn, 2009; this study also determined that the most 5' bovine TCRVα gene on chromosome 10 was located at 24.94Mb.

Comparison of the incomplete bovine assembly to the corresponding region of the human genome (Chromosome 14q11.2) revealed a similar organisation. The bovine Cα region, which mapped to chromosome 10 at 22.59Mb, corresponded with the human 'IG C gene Ig/T-cell receptor' gene on chromosome 14. 0.04Mb downstream from this on chromosome 14 were human 'IG J gene Ig/T-cell receptor' genes, which corresponded to the location of bovine Jα33 on *Bos taurus* chromosome 10 at 22.63Mb, amongst other genes identified as "Ig-like". The group of identified bovine "Ig-like" V genes at 24.4 – 24.9Mb occupied a corresponding position to a large number of human 'IG V gene Ig/T-cell receptor' genes (see figure 3-18). At the extreme downstream end of this group are the human TCRVα7.2 (AE000658.6) and TCRVα7.1 (AE000658.4) genes, identified during genome construction from the IMGT database. These two genes are bordered by olfactory genes: OR10G2 (between Vα7.1 and Vα7.2), OR4E2 (between Vα7.2 and subsequent TCR Vα genes), and OR4E1P (olfactory receptor psuedogene), respectively, which occupy very similar positions to the olfactory genes identified on bovine chromosome 10. These available data indicate that the respective genes in humans and cattle occupy similar positions within the alpha locus.

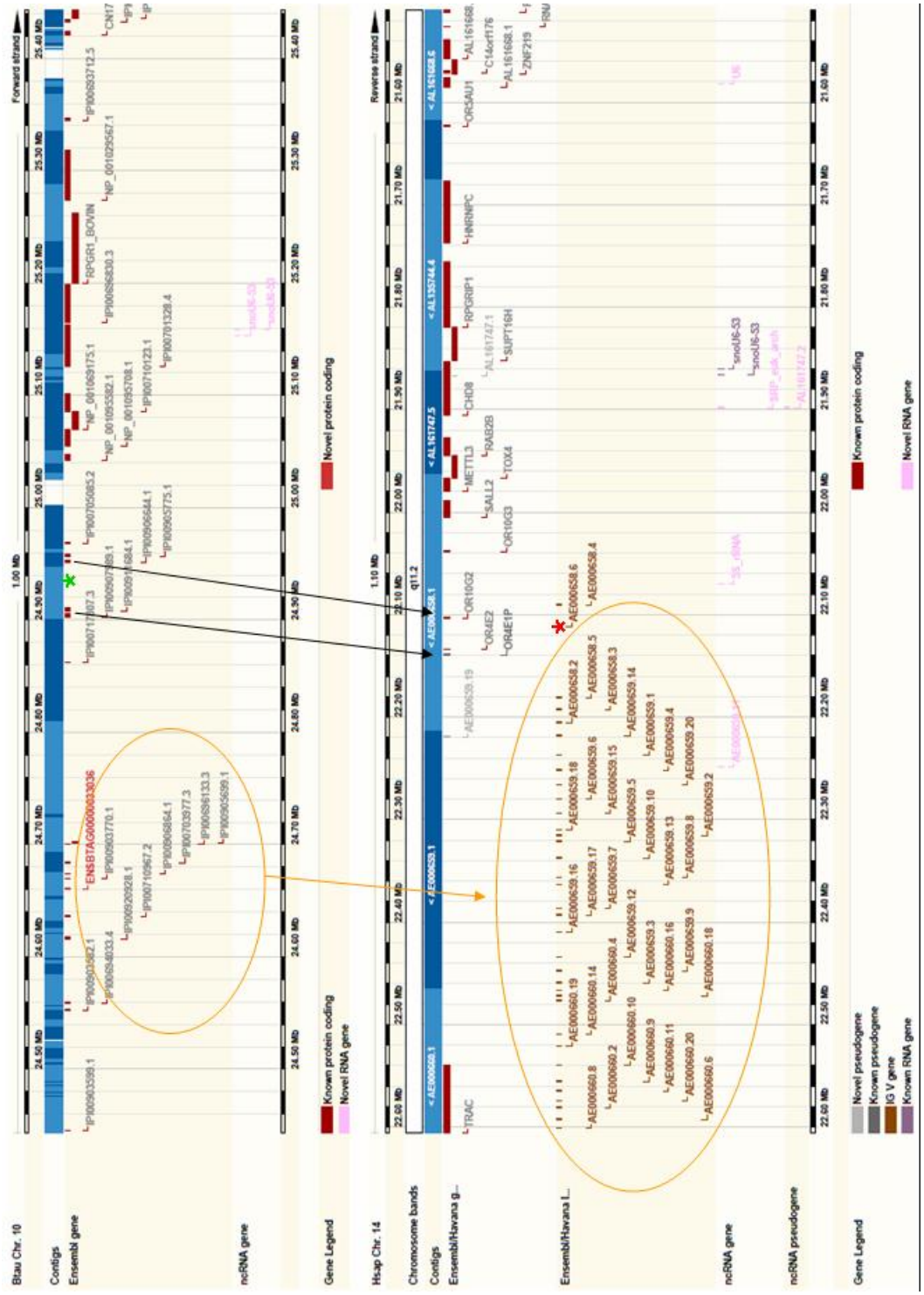


Figure 3-18: Partial *Bos taurus* genome assembly showing BLAST location of Vα19 (*), and comparison to the corresponding region of the *Homo sapiens* genome. Human Vα7.2 (*) is indicated. The corresponding neighbouring olfactory genes (black arrows) and other TCRVα genes (orange rings) suggest that bovine Vα19 occupies a position at the 5' extremity of the TCRα locus like its human orthologue.

3.4 Discussion

Human and murine MAIT cells express an invariant TCRα chain incorporating the orthologous V and J gene segments and with near identical CDR3 region sequences. A partial TCRα sequence had been previously identified as the orthologous Vα gene of MAIT cells in cattle (Tilloy et al. 1999), but full-length bovine sequence had not been published. Additionally, searches on online databases showed that no orthologous gene had been recognised in sheep. Using the partial bovine sequence as a template, the results in this chapter reveal the full-length MAIT TCRα sequences for both cattle and sheep. This confirms presence of ovine MAIT cells, and extends the previous information on MAIT TCRα sequences.

These new sequence data confirm that there is a very high level of conservation of the MAIT TCRα chain between sheep, cattle, humans and mice. The MAIT transcripts utilise the orthologous V and J gene segments in the different species. Conservation of the corresponding V genes is especially high between sheep and cattle, at 93%. Importantly, the CDR1 region is completely conserved between all species. The CDR2 region of sheep, cattle and humans is also identical, apart from one codon (position one of five) of the murine CDR2 which differs from other species. This change, which involves a substitution of the polar amino acid, asparagine, with the non-polar amino acid valine, is potentially significant; the hydropathy index of these molecules is widely different (+4.2 compared to -3.5) and this substitution may potentially affect the overall nature of the bovine MAIT CDR2 loop in association with water or lipid, respectively.

The MAIT CDR3 region is also highly conserved between these species. CDR3 length in each is identical, at 8 amino acids long, which is somewhat shorter than the average TCRα CDR3 length of 10 residues (Moss and Bell 1995). A few conservative substitutions are observed in the CDR3 sequence between species. As previously identified, the codon in position one is different for human, mouse and bovine sequences, but the human and murine CDR3 sequences are otherwise identical. The bovine CDR3 sequence contains different amino acids at positions 3

and 7, glycine for serine (both weakly hydrophilic) and tryptophan for leucine, respectively. Although tryptophan is weakly non-polar, leucine is a strongly polar amino acid and hydrophobic, and it is arguable that this substitution may affect the conformation of the CDR3 loop. The newly identified ovine MAIT CDR3 region only differed from the human and murine sequences at position 6, where it contained arginine rather than glutamine. Arginine is strongly basic whereas glutamine is neutral, but both of these amino acids are strongly hydrophilic. In summary, the cross-species conservation in the use of orthologous V and J gene segments and the high degree of similarity in the sequences of all 3 CDR segments between species strongly indicate that both the restriction element and the putative ligand(s) of the MAIT TCR are also conserved between species.

In studying MAIT sequences from cloned DNA, a variation in the first CDR3 codon within animals was noted in roughly 1 in 4 bovine and ovine clones. These clones contained a single base pair substitution, A for G, which altered the coding of the amino acid at position 1 of the CDR3 from methionine to isoleucine. These amino acids have similar properties, being non-polar and neutral, and hence this substitution is unlikely to result in a major structural change in the CDR3 loop. This variation in cDNA sequence can be accounted for by differential nucleotide deletions at the VJ junction. Human and murine genomic Vα19 and Jα33 segments are the same length as the bovine genes; hence, their generation of the VJ junctions to form the short MAIT CDR3 (8 amino acids) occurs primarily by nucleotide deletions. In the mouse, a short homologous sequence at the 3' end of Vα19 and 5' end of Jα33 results in a predominant canonical nucleotide sequence which codes for AVRDS over the VJ junction. This sequence was detected in 80 of 101 murine MAIT clones (Treiner et al. 2005); other clones encoded the AVMDS sequence (caused by alternate deletion), or other sequences of identical length but with some amino acid variability. In contrast, human genomic Vα19 and Jα33 sequences differ at their respective 3' and 5' ends, which results in some variability of the two N-terminal CDR3 amino acids of the human MAIT Vα; these include AVMDS and AVRDS, 'germline' sequences involving deletions but without 'N-region' additions, as well as AVKDS and AAXDS (where X is M, R, K, or L), formed by nucleotide additions.

However, the CDR3 region was always of constant length (Treiner et al. 2005). Variations in the sequence of the MAIT TCRα CDR3 may allow the TCR to recognise slightly different ligands, or may simply suggest that these sequences result in TCRs with very similar structure and specificity (Porcelli et al. 1993; Tilloy et al. 1999; Illes et al. 2004; Treiner et al. 2005). Alternatively, variations in the MAIT CDR3 region may relate to providing flexibility in combining with different TCRβ chains. Human and murine MAIT cells have been shown to exhibit an oligoclonal TCRβ repertoire, consistent with peripheral expansion (Tilloy et al. 1999). These cells preferentially express particular Vβ genes: Vβ2 and 13 in humans, and Vβ6 and 8 in mice, where murine Vβ6 and 8 represent the murine orthologue of human Vβ13. Treiner et al, 2005, hypothesised that this Vβ repertoire was driven by the selecting antigen, since Vα19-Jα33 transgenic mice showed an increased frequency of $V\beta8^+$ and $V\beta6^+$ T-cells, which reverted to control levels when the transgenic mice were crossed onto an MR1 KO background (therefore preventing MAIT cell selection). However, the overall TCRβ diversity of each species' MAIT population remains extensive, since most Jβ genes and 7 (human) / 5 (murine) Vβ segments have been demonstrated on MAIT cells, in association with CDR3 of variable lengths (Porcelli et al. 1993; Tilloy et al. 1999; Treiner et al. 2005).

In humans and mice, MAIT cells are highly enriched in the $CD4^-$ $CD8^-$ T-cell population (Tilloy et al. 1999). Preliminary studies conducted by these authors also suggested that they are enriched in this population in cattle. However, analyses of sequences from sorted subsets of bovine T-cells suggest that MAIT cells are neither enriched nor confined to this subset in cattle. Although the numbers of sequences studied were relatively small, any substantial enrichment within this small subset, which was shown to represent up to 4% of T cells in cattle, should have been evident from these analyses. One possibility is that MAIT cells may instead be enriched in a different subset of bovine T-cells. Similarly, although the data achieved was not quantitative, the detection of 4 MAIT sequences out of a total of 89 examined in PBMC indicate that MAIT sequences are present at a much higher frequency than would be expected as a result of random rearrangement, in the absence of further

expansion of cells bearing the rearrangement. Studies to estimate the frequencies of MAIT cells are described in chapter 4.

The V and J segments used by the MAIT TCRα were detected in combination with other J and V segments, respectively. However, none of these rearrangements was detected more than once; hence, there is no evidence that these cell-populations had undergone expansion. Although some had CDR3 region sequences similar to the MAIT sequence, by virtue of sharing the J segment, the CDR1 and CDR2 sequences in the associated V genes differed substantially from those of the MAIT V gene. These findings demonstrate that the Vα19 and Jα33 segments are not used exclusively for generating MAIT TCRα chains in cattle. Therefore, PCR assays specific for these individual gene segments cannot be used reliably to detect and quantify MAIT cells.

Analyses of the bovine genome sequence database suggested that, as in humans and mice, the bovine MAIT Vα19 segment is located at the 5' end of the TCRα locus. Although assembly of the locus is incomplete, a large number of Vα genes are identified 3' of Vα19 but none are identifiable on the 5' side. The Vα19 gene is also bordered by olfactory receptor genes similar in sequence and in relative location to olfactory receptor genes found on either side of the human Vα7.2 gene. The process of rearrangement of TCRα V and J gene segments initially involves those V gene segments proximal to the J genes and progresses to those at the 5' end of the locus (Treiner et al. 2005). Since the MAIT Vα gene is the most 5' V gene segment, there is little obvious potential for preferential rearrangement. If rearrangement is random, the large number of Vα genes (estimated to be >300) within the bovine TCRαδ locus, coupled with multiple J gene segments (Reinink and Van Rhijn 2009), even in the absence of junctional diversity would result in a very low frequency of T cells with the Vα19 and the Jα33 gene segments. This implies that there must be clonal expansion of cells expressing the MAIT TCRα chain. Theoretically, expression of the restricting MHC together with a 'conserved' self ligand within the thymus could give rise to positive selection and expansion of the MAIT cells in the thymus.

However, recent studies indicate that expansion occurs peripherally following their exit from the thymus (discussed in more detail in chapter 4).

In conclusion, ruminant MAIT TCRα sequences show a high degree of sequence similarity to those described for humans and mice. This conservation is particularly apparent in the CDR region sequences, and in the case of CDR3 indicates that there is selective conservation of the sequences generated by junctional diversity during the process of TCRα rearrangement. Analysis of TCRα sequences from cDNA indicate that, as in other species, cells expressing these TCRα chains are present at a much higher frequency than would be expected from random rearrangement, indicating that they have undergone multiplication following thymic development.

4 Frequency and Distribution of MAIT cell TCRVα Expression

4.1 Introduction

MAIT cells were originally defined in humans and mice by an invariant TCRα chain, which has a CDR3 of constant length, and is highly conserved between species. Results presented in chapter 3 demonstrated that MAIT TCRα in sheep and cattle are similarly invariant and highly conserved. The MAIT TCRα is detected in peripheral blood at much higher frequencies than would be expected from random TCR rearrangement, and hence it has been suggested that they might be capable of responding rapidly to appropriate immune stimuli. The conserved nature of the TCR may also allow it to recognise and respond rapidly to conserved antigen/pathogenic patterns, in a manner similar to the PRRs of the innate immune system.

Although some phenotypic data are available for MAIT cells in humans and mice, there are no definitive surface markers that can be used to distinguish them from other T-cells. Analyses of their frequencies have therefore relied on the use of qPCR assays to measure MAIT TCRα in relation to total TCRα in cDNA prepared from cell populations of interest.

Initial studies using these assays in humans and mice indicated that the frequency of MAIT cells is 5-10 times higher in humans than in mice (Treiner et al. 2005). In comparison to peripheral blood T-cells, MAIT TCRα expression was enriched within the $CD4^-$ $CD8^-$ T-cell populations of both species; human MAIT cells accounted for up to 15% of these cells. In contrast, MAIT cells comprised only ~2% of murine lymph node $CD4^-$ $CD8^-$ T-cells. Overall, between 0.15 and 0.2% of all human peripheral blood T-cells were estimated to express the MAIT TCRα; however, a definitive figure for murine MAIT TCRα expression in peripheral blood T-cells has not been published (Tilloy et al. 1999).

Initial studies on MAIT cells in mice suggested that their development was dependent on the thymus and bacterial intestinal flora, as MAIT cells were absent from *nude*, athymic mice (Tilloy et al. 1999), and germ-free mice (Treiner et al. 2003). MAIT cell development was also shown to be dependent on B-cells, as MAIT cells were absent from both mice and people deficient in B-cells (Treiner et al. 2003). However, these early studies failed to detect MAIT cells by qPCR in human or murine thymus; they also could not be detected in the spleen of neonatal mice. MAIT TCRα transcripts were detected in peripheral blood of young children, but this was at a lower level than that detected in adults (Tilloy et al. 1999). A recent study has now shown that MAIT cells with a naïve T-cell phenotype are present at very low frequency in human thymus and cord blood, whereas they are more abundant and have a memory phenotype in adult peripheral blood (Martin et al. 2009). These recent findings, together with other data, suggest that MAIT cell development involves an initial thymic selection which is dependent on MR1 expression, followed by a population expansion in the periphery which is dependent on B-cells and commensal gut flora. In contrast, this study demonstrated that murine MAIT cells remain in low number with a naïve phenotype in the periphery, even in the presence of B-cells and gut flora, suggesting possible absence of clonal expansion in this species. However, B-cells were shown to play a role in the accumulation of MAIT cells in the mesenteric lymph nodes of experimental mice, and experimental adoptive transfer of $MR1^{+}$ B-cells into mice with a naïve MAIT cell population was sufficient to promote MAIT cell accumulation with acquisition of a memory phenotype (Martin et al. 2009).

Analyses of MAIT expression in tissues from humans and mice have revealed differences in distribution in peripheral tissues. MAIT cells are particularly abundant in the intestinal mucosa (lamina propria lymphocytes), from whence they derive their name, and the mesenteric lymph nodes, in comparison to intra-epithelial and peripheral blood lymphocytes. Murine MAIT transcripts were shown to be ten-fold more abundant in intestinal mucosa than in mesenteric nodes. MAIT cells are also present within lung tissue, which suggests that they may also accumulate in other mucosal tissues (Treiner et al. 2003).

Initial analyses of lymphocyte subsets defined by expression of CD4 or CD8 indicated that MAIT cells are enriched within the small subset of $CD4^{-}CD8^{-}$ αβ T-cells in comparison with the $CD4^{+}$ and $CD8^{+}$ populations (Tilloy et al. 1999). However, they were also detected at lower frequencies in other subsets, including human $CD8\alpha\alpha^{+}$ T-cells and in murine $CD4^{+}$ T-cells. More recent publications have suggested that MAIT expression is also enriched in the human $CD8\alpha\beta^{+}$ population (Martin et al. 2009). In the initial report of a partial bovine MAIT TCRα sequence, MAIT transcripts were also apparently enriched in the $CD4^{-}CD8^{-}$ population in bovine peripheral blood (Tilloy et al. 1999). However, results presented in chapter 3 of this thesis suggest that this might not be the case.

The work described in this chapter aimed to develop a specific qPCR assay that could be used to provide a quantitative estimate of the numbers of MAIT cells within T-cell populations of cattle and sheep. This assay was used to determine the distribution of MAIT cells in different tissues and within different T-cell populations, and to examine the effect of age on their frequency.

4.2 Materials and Methods

4.2.1 Experimental animals

Blood samples for isolation of PBMC were collected from adult healthy Holstein-Friesian cattle, as described in section 2.1.1. Tissue and blood samples for qPCR analysis of MAIT cell distribution were taken from two 6-month old Blackface-cross ewe lambs and from Holstein-Friesian calves, which were sampled either within 24 hours of birth (neonatal), or at 3-weeks or 3-months old. Blood was collected immediately ante-mortem (section 2.1.1).

4.2.2 Post-mortem sample collection and RNA extraction

Tissue samples were collected immediately post-mortem using sterile instruments (see section 2.1.2), and trimmed to fragments of <5mm thickness before storage in RNAltr (Applied Biosystems, Warrington, UK). Tissue homogenisation and RNA extraction are described in section 2.4.1.2; RNA was extracted from PBMC using Tri-Reagent (section 2.4.1.1). RNA extracted from PBMC and tissue samples for qPCR was treated with DNase before reverse transcription (sections 2.4.1.4; 2.4.2.1).

4.2.3 Real-time quantitative PCR

The qPCR protocol is described in section 2.4.2.4. qPCR primer sequences are shown in table 4-1. Primers "I"/"Jr" (bovine) and "boovUntr3f"/"CorovC3r" (ovine) were used to generate cDNA clones containing bovine/ovine MAIT TCRα chain sequence (each consisting of full-Vα and partial-Cα regions; for these primer sequences, see section 3.2.3). DNA from plasmids containing these clones was used as a positive control and to generate qPCR standard curves. The MAIT-cell V-region-specific qPCR extended across the V-J junction, utilising the "boovAV19f" forward primer (cattle and sheep), located within the V gene segment, with reverse primers "boovAJ33r" (bovine) or "ovAJ33r" (ovine) located in the J gene. A 33bp

probe: "boAV19" lying near the 3' end of Vα19 was used. For the α chain constant-region-specific qPCR, primers "NGconst-f" and "NGconst-r", and probe "NGconst", were used for both cattle and sheep. 3µg of cDNA was used for each reaction in the described protocol. All qPCR samples for optimisation and experimental use were run in triplicate, both for standards and unknowns. Average results with standard errors were used for analysis of results. The standard error (SE) of relative abundance of MAIT cells (MAIT transcript (x) as a percentage of overall αβ TCR transcript (y)) was calculated using the formula:

$$SE\ (x/y) = Sq\ Rt\ ((SE(x)/(y))^2+(((x)*SE(y))/(y)^2)^2).$$

The qPCR reaction efficiency was calculated using the formula: $E = 10^{(-1/slope)}-1$.

qPCR data for two 3-month old calves was provided by Dr Ildiko van Rhijn, University of Utrecht, using the same methods (see section 2.4.2.4).

Table 4-1: Primer and probe sequences used for qPCR analysis.

Target Sequence	**Forward** **5' sequence 3'**	**Taqman Probe** **5' sequence 3'**	**Reverse** **5' sequence 3'**
MAIT-cell V-region (bovine)	boovAV19f CATTCCTTAGACGCTCTG ATGCACA	boAV19 FAM-TTACCTCCTTCTGAAGGAACT CCACATGAAAGA-BHQ	boovAJ33r GCAACTATCAGTGGATCTGG GGC
MAIT-cell V-region (ovine)	boovAV19f -	boAV19 -	ovAJ33r GCAACTATCGGTTGATCTGG GGC
TCRα chain C-region	NGconst-f ACGGGATAGTGACTTGGG GAA	NGconst FAM-ACACCTTCAACGAGAACA-BHQ	NGconst-r CCAACGTGGTCGAGAAAAGC TTTGAAA

On Taqman probes, FAM: fluorescent dye; BHQ: quencher.

4.2.4 Purification and PCR of T-cell subpopulations

Immunofluorescence staining and analysis of cells by flow cytometry are detailed in section 2.3.3. Monoclonal antibodies specific for CD3 (MM1A), CD4 (IL-A12), CD8 (IL-A51) (IL-A105), and γδ TCR (GB21A) were used for cell staining. In one experiment, cells were stained with a mixture of MM1A, IL-A12, IL-A105 and GB21A antibodies and purified by cell sorting into $CD3^+$ $CD4^+$ $CD8^+$ $\gamma\delta^+$ and $CD3^+$ $CD4^-$ $CD8^-$ $\gamma\delta^-$ populations. In a second experiment, cells were also stained with individual antibodies specific for CD3 (MM1A), CD4 (IL-A12) and CD8 (IL-A51) and sorted into positive and negative populations for each marker. Purified cell fractions were re-analysed and found to be 98-99% pure.

PCR assays were performed using reaction set-up and reagents as described in section 2.4.2.2. A standard reaction was used, consisting of 10 minutes at 95°C, 35 cycles of (1 minute at 95°C, 1 minute at a 60°C, 1 minute at 72°C), and a final extension period of 10 minutes at 72°C. Primers sequences for "boovAV19f"/"boovAJ33r" and "G3PDH1/2" are provided in sections 3.2.3 and 2.4.3.2, respectively. PCR products were analysed on 1.5% agarose gels, as described in section 2.4.3.3.

4.2.5 Efferent gastric lymph samples from sheep undergoing experimental *Telodorsagia circumcincta* infection

Ovine gastric lymph samples were provided by Dr David Smith at the The Moredun Research Institute, Edinburgh. These samples had been collected following surgical cannulation of the efferent lymphatic draining the gastric lymph node of sheep undergoing primary or secondary infections with the abomasal parasite *Telodorsagia circumcincta* (see section 2.1.2). Animals undergoing secondary infection had been previously subjected to a trickle infection with *Telodorsagia circumcincta*. Both groups of animals were treated with anthelmintic 7 days prior to challenge. The sheep were challenged with 50,000 L3 larvae 1-7 days after cannulation. Samples

were provided as total RNA; these were processed for qPCR as for previous samples (section 4.2.3).

4.2.6 Statistical analyses

Real-time PCR results from three neonatal, three 3-week old, and three 3-month old calves were assessed using general linear (mixed effect) models on "R" version 2.9, a programming language and software environment for statistical computing, available from http://www.r-project.org/.

4.3 Results

4.3.1 Design, standards and optimisation of 'Real-time' qPCR

'Real-time' quantitative PCR assays were designed to measure expression of the MAIT TCRα chain in relation to total TCRα chains, to determine the relative abundance of MAIT cells within the overall αβ T-cell population (see figure 4-1). The MAIT TCRα PCR utilised a forward primer within the V gene segment and a reverse primer in the MAIT CDR3 of the J gene, covering 17/24bp of 3' CDR3 sequence and additional J segment sequence (see figure 4-2). The pan-TCRα PCR used forward and reverse primers within the C gene segment. The specificity of the primer pairs was tested by cloning the PCR products obtained from cDNA prepared from PBMC and sequencing 4 clones from each: all of four clones for each reaction matched the target sequence.

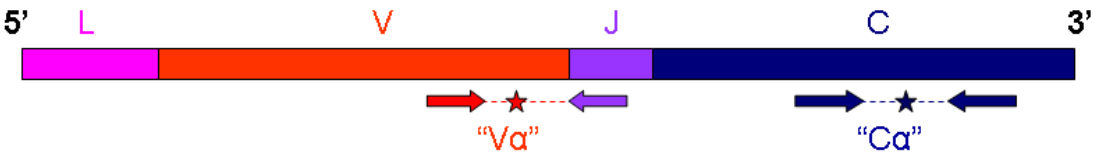


Figure 4-1: Location of qPCR primers (arrows) and fluorescent probes (stars) used to detect frequency of MAIT TCRα expression: The "Vα" reaction used a forward primer specific for bovine/ovine Vα19, a probe specific for 3' Vα19, and a reverse primer specific for the MAIT CDR3 region in Jα33. The second "Cα" reaction used forward and reverse primers, and a probe, specific for the constant gene segment. L: Leader; V: Vα19 gene; J:Jα33 gene; C: Constant gene.

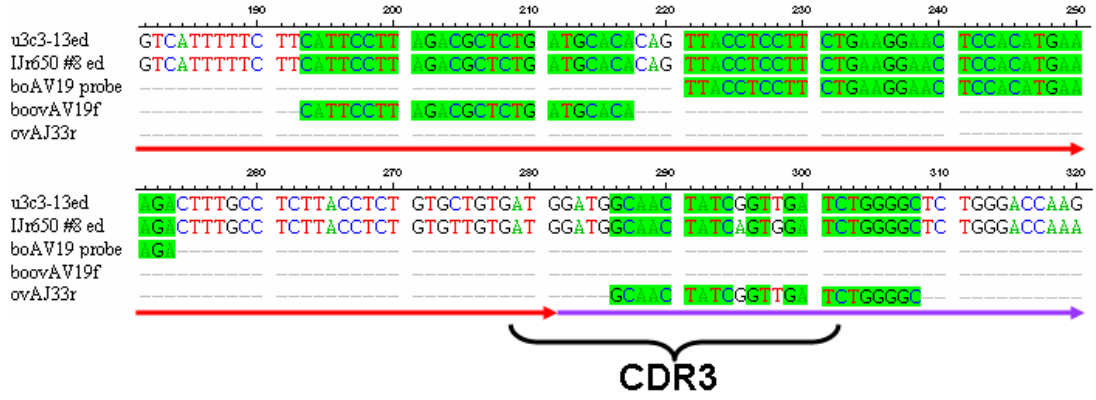


Figure 4-2: V-J junctions of MAIT TCRα sequences (ovine: u3c3-13ed) (bovine: IJr650#8ed) used as standards for qPCR. Vα19 sequence is indicated with a red arrow, and Jα33 sequence with a purple arrow; the CDR3 is as marked. The location of "Vα" reaction primers (boovAV19f, position 193-217, and ovAJ33r, position 286-308) and probe (boAV19, position 221-253) are highlighted in green. The forward primer and probe sequences were exact matches for both ovine and bovine sequence, and were used for assays of cDNA from both species. There were two nucleotide differences between ovine and bovine sequence within the site of the reverse primer – these are un-highlighted at 295 and 298 within the highlighted reverse primer location – hence a different specific reverse primer was used for each species (ovAJ33r and boovAJ33r, respectively). Primer boovAJ33r shared exact location and length with ovAJ33r but contained 2 single nucleotide differences from ovAJ33r to match differences in bovine sequence from ovine. These reverse primers included the 3' 17bp of MAIT CDR3 sequence.

Bovine and ovine standards for qPCR were created from plasmid DNA extracted from fully sequenced MAIT TCRα chain cDNAs (see section 3.3.1). Serial dilution of standards was used to establish optimal primer and probe concentrations for qPCR reactions, as well as optimal annealing temperature and thermal-cycling conditions, which are detailed in section 2.4.2.4. Primers were tested using a concentration matrix at two different annealing temperatures (Tm), and optimal combinations were chosen based on speed of amplification (low Ct), accuracy (low standard error (S.E.) between results), and change in fluorescence during reaction (high ΔRn) – this was 53°C and 500nM for all four primers, see table 4-2, A and B. Subsequently, probes

were tested at four concentrations for the same parameters; both probes functioned best at 250nM, see table 4-2, C.

Table 4-2: qPCR Optimisation.

A) Results of qPCR optimisation testing primers boovAV19f/boovAJ33r on MAIT TCRα plasmid DNA in a matrix of 50nM, 300nM, 500nM and 900nM concentrations at different annealing temperatures (Tm), 53°C and 60°C. Primer concentrations in the first column are denoted for forward / reverse primers. The selected conditions for use (Tm:53°C, 500nM/500nM) are highlighted red. Ct: crossing threshold; S.E.: standard error between results performed in triplicate; ΔRn: delta reaction (change in fluorescence during reaction).

B) Results of qPCR optimisation of primers NGconst-f/NGconst-r, as for table A. The selected conditions for use (Tm:53°C, 500nM/500nM) are highlighted blue.

C) Results of qPCR optimisation of probes boAV19 and NGconst, tested at a range of concentrations, using primer concentrations 500nM/500nM and Tm:53°C. The selected conditions for use are highlighted red and blue, respectively.

Table 4-2, A) qPCR boovAV19f / boovAJ33r primer optimisation

Primer Conc'n	Tm:53°C			Tm:60°C		
	Ct	S.E.	Δ Rn	Ct	S.E.	Δ Rn
50/50	35.08	0.121	1.31	41.27	0.622	0.52
50/300	31.66	0.248	2.50	36.80	0.341	1.39
50/500	31.20	0.085	2.70	35.67	0.432	1.63
50/900	30.59	0.241	2.88	33.75	0.337	1.92
300/50	29.63	0.253	1.27	32.10	0.309	0.79
300/300	27.11	0.098	2.85	28.77	0.259	2.10
300/500	26.76	0.088	3.04	28.09	0.195	2.54
300/900	26.32	0.090	3.38	26.96	0.204	2.76
500/50	29.71	0.131	1.22	31.75	0.415	0.81
500/300	27.07	0.041	2.64	28.42	0.082	1.85
500/500	26.82	0.082	3.13	27.45	0.332	2.53
500/900	26.46	0.136	3.32	26.52	0.112	2.88
900/50	29.84	0.258	1.14	32.49	0.682	0.71
900/300	27.19	0.098	2.34	28.35	0.047	1.64
900/500	26.98	0.019	2.81	27.96	0.217	2.04
900/900	26.74	0.185	2.96	26.41	0.264	2.86
NTC	-	-	-	-	-	-

Table 4-2, B) qPCR NGconst-f / NGconst-r primer optimisation

Primer Conc'n	Tm:53°C			Tm:60°C		
	Ct	S.E.	Δ Rn	Ct	S.E.	Δ Rn
50/50	27.84	0.187	1.33	26.90	0.089	0.97
50/300	26.07	0.024	2.31	24.34	0.087	1.89
50/500	25.46	0.068	2.69	24.03	0.056	2.15
50/900	25.35	0.145	2.59	23.81	0.437	2.18
300/50	28.29	0.218	1.50	27.73	0.216	1.19
300/300	25.92	0.101	2.88	24.86	0.091	2.15
300/500	25.64	0.042	3.12	24.51	0.088	2.54
300/900	25.15	0.019	3.50	24.23	0.159	2.84
500/50	28.10	0.097	1.79	27.78	0.108	1.37
500/300	26.07	0.139	2.91	25.13	0.054	2.17
500/500	25.62	0.055	3.47	24.76	0.008	2.80
500/900	25.41	0.245	2.95	24.68	0.066	2.57
900/50	28.45	0.182	1.57	28.35	0.184	1.18
900/300	26.53	0.098	2.78	25.44	0.357	2.47
900/500	25.81	0.046	3.45	25.10	0.018	2.82
900/900	25.43	0.018	3.46	24.82	0.060	2.90
NTC	-	-	-	-	-	-

Table 4-2, C) qPCR boAV19 and NGConst probe optimisation

	boAV19			NGConst		
Probe Conc'n	Ct	S.E.	Δ Rn	Ct	S.E.	Δ Rn
50nM	28.02	0.016	0.79	26.97	0.030	0.57
100nM	27.20	0.045	1.57	25.73	0.094	1.68
150nM	26.99	0.079	2.35	26.16	0.068	1.11
250nM	26.70	0.146	3.49	24.87	0.047	4.32

The reaction efficiency was calculated for both standard curves of each qPCR assay performed, using the equation $E = 10^{(-1/slope)}-1$. Only results from qPCR assays with reaction efficiencies between 90-110% were used for further analyses. Table 4-3 shows the reaction efficiencies of each qPCR assay used in this study.

Table 4-3: Reaction efficiency of each qPCR assay performed.

	Reaction Efficiency	
qPCR Experiment	**Vα**	**Cα**
Calf 2511	100.1%	94.3%
Calf 0312	100.0%	91.8%
Calf 0412	101.0%	96.1%
Calf 1776	91.3%	92.5%
Calf 1777	97.4%	92.0%
Calf 1803	95.7%	93.2%
Calf 2703	95.0%	90.4%
Sheep 3156E – run#1	92.0%	95.0%
Sheep 3156E – run#2	101.2%	95.0%
Sheep 3215E – run#1	90.5%	100.0%
Sheep 3215E – run#2	104.7%	98.1%
Adult bovine PBMC	90.3%	91.9%
T-cell subsets: CD3 vs CD4/CD8/γδ	98.0%	107.7%
T-cell subsets: CD3, CD4, CD8	95.9%	92.5%
Sheep B092 (*T. circ*)	99.1%	94.0%
Sheep B067 (*T. circ*)	94.0%	100.8%
Sheep 1063 (*T. circ*)	95.8%	98.9%
Sheep B103 (*T. circ*)	95.5%	98.3%
Sheep 992 (*T. circ*)	96.8%	95.1%
Sheep 955 (*T. circ*)	93.5%	98.6%
Sheep B189 (*T. circ*)	100.8%	101.6%
Sheep B161 (*T. circ*)	107.8%	97.2%
Sheep B1465 (*T. circ*)	100.5%	100.3%
Sheep B1408 (*T. circ*)	102.9%	98.0%
Sheep B1399 (*T. circ*)	108.6%	99.4%

Reaction efficiencies were calculated using the formula $E=10^{(-1/slope)}-1$. Each qPCR experiment for an individual animal relates to all blood and tissue samples from that animal. 'Adult bovine PBMC' relates to peripheral blood mononuclear cells from four adult cattle: #109, #158, #1706, #1707. 'T-cell subsets: CD3 vs CD4/CD8/γδ' relates to purified bovine T-cell sub-populations of $CD4^+$+$CD8^+$+$γδ^+$ cells, $CD3^+CD4^-CD8^-γδ^-$ cells, and unfractionated PBMC. 'T-cell subsets: CD3, CD4, CD8' relates to bovine T-cell sub-populations purified as positive or negative for each phenotype (CD3, CD4, CD8). 'Sheep xxxx (*T. circ*)' relate to efferent lymph samples collected from sheep experimentally infected with *T. circumcincta*.

4.3.2 Levels of expression of bovine MAIT TCRα are highest in spleen and ileum, and increase after the neonatal period

The optimised qPCR assay was used to measure MAIT TCRα transcripts as a means of assessing MAIT cell prevalence and distribution in cattle. Neonatal and juvenile calves were sampled to assess the effect of age upon MAIT cell levels and location. PBMC and samples from intestinal and immunological tissues (thymus, lymph node

(prescapular and mesenteric), spleen, small and large intestinal walls) were collected from three neonatal calves (within 24 hours of birth), three 3-week old calves, and one 3-month old calf. The results from animals of different ages were first considered together. The levels of expression of MAIT TCRα in comparison to the total TCR α pool ranged from 0.1% to 1.8%, depending on tissue and age. Highest levels of MAIT transcripts were detected in spleen and mesenteric lymph node (up to 1.8% and 0.9%, respectively). All other tissues showed comparatively lower MAIT transcript levels; intermediate expression was seen in small and large intestinal walls and peripheral lymph node, while thymus and PBMC showed very low expression (see figure 4-3). Statistical analyses of data from all age groups for each tissue are shown in table 4-4, and showed MAIT transcript levels in the spleen to be highly significantly raised when compared to thymus ($p<0.001$), PBMC ($p<0.001$) and peripheral lymph node ($p<0.001$). Percentages of MAIT transcript in mesenteric lymph node, ileum and colon were highly significantly increased in comparison to thymus ($p<0.001$, $p<0.001$, $p=0.002$) and significantly increased in comparison to PBMC ($p=0.016$, $p=0.015$, $p=0.041$).

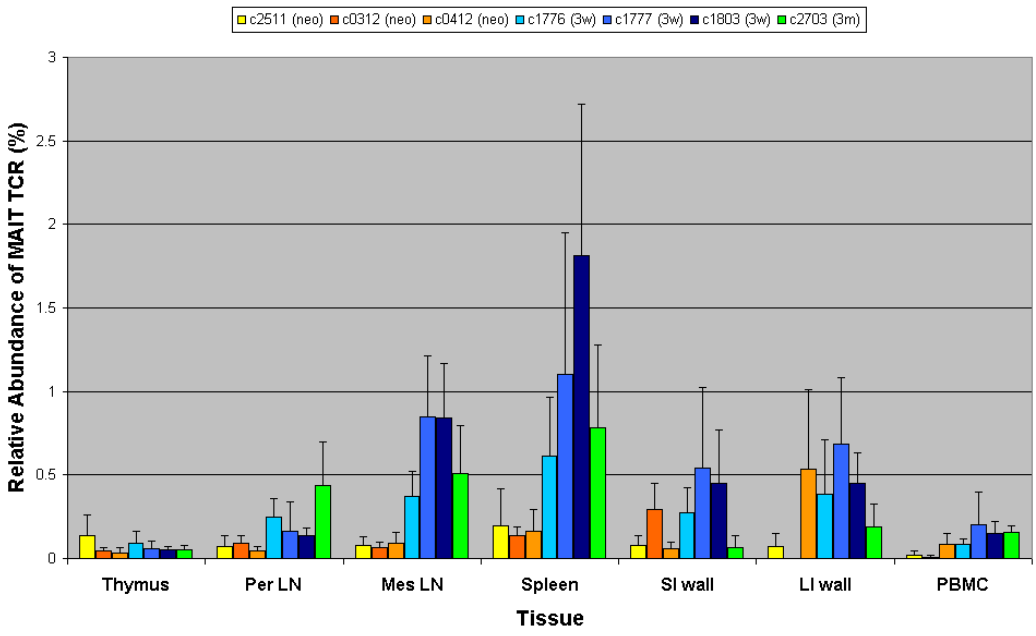


Figure 4-3: Relative abundance of MAIT TCRα expression in different bovine tissues. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

Table 4-4: Results of statistical analyses of relative abundanceof MAIT TCRα transcripts in different tissues from calves of varying ages.

Tissues compared	p value
Per LN – Thymus	0.10049
Mes LN – Thymus	< 0.001 ***
Spleen – Thymus	< 0.001 ***
SI wall – Thymus	< 0.001 ***
LI wall – Thymus	0.00248 **
PBMC – Thymus	0.98084
Mes LN – Per LN	0.68049
Spleen – Per LN	0.00105 **
SI wall – Per LN	0.75751
LI wall – Per LN	0.88183
PBMC – Per LN	0.51266

Tissues compared	p value
Spleen – Mes LN	0.33190
SI wall – Mes LN	0.99999
LI wall – Mes LN	0.99965
PBMC – Mes LN	0.01567 *
SI wall – Spleen	0.14216
LI wall – Spleen	0.10846
PBMC – Spleen	< 0.001 ***
LI wall – SI wall	1.00000
PBMC – SI wall	0.01495 *
PBMC – LI wall	0.04051 *

***** = very highly significant, ** = highly significant, * = significant. Tissues included: thymus, peripheral lymph node (Per LN), mesenteric lymph node (Mes LN), spleen, small intestinal wall (SI wall), large intestinal wall (LIwall), peripheral blood mononuclear cells (PBMC).**

The effect of age upon MAIT cell prevalence and distribution was examined by comparing the results obtained from the 3 different age groups. Additional results for two 3-month old calves were included in the analysis (although the mesenteric lymph node was not tested in these animals); results are presented in figure 4-4. Neonatal animals showed consistently low levels of MAIT TCRα expression throughout all tissues tested, and statistical analysis of experimental data showed a highly significant increase in MAIT transcript abundance in 3-week ($p<0.001$) and 3-month ($p=0.044$) animals when compared to neonates (see table 4-5). There was no significant change in MAIT transcript abundance between 3-week and 3-month animals, suggesting that most of the expansion of the MAIT population has occurred by 3 weeks of age.

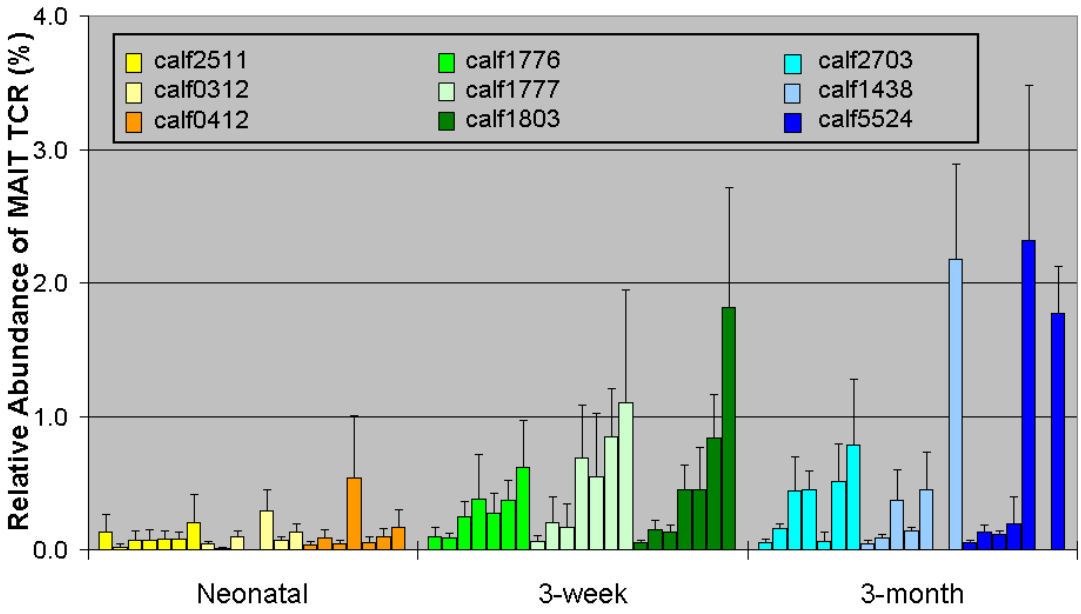


Figure 4-4: Relative abundance of bovine MAIT TCRα expression in tissues from three animals within each of three age groups: neonatal (2511, 0312, 0412), 3-week old (1776, 1777, 1803), and 3-month old (2703, 1438, 5524). Individual tissue results for each animal are shown in the order: thymus, peripheral blood mononuclear cells, peripheral lymph node, large intestinal wall, small intestinal wall, mesenteric lymph node, spleen. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

Table 4-5: Results of statistical analyses of relative abundance of MAIT TCRα transcripts in tissues from calves in three different age groups – neonatal, 3-week old and 3-month old.

Age groups compared	p value
3week – Neonatal	< 0.001 ***
3month – Neonatal	0.0044 **
3month – 3week	0.7315

When both individual organ and age of animal are taken into account, spleen and mesenteric lymph node show marked increases in MAIT expression by 3-weeks of age, while ileum, colon, peripheral lymph node and PBMC show less pronounced

increases (see figure 4-5). MAIT expression in the thymus is consistently low in all ages. Changes in MAIT expression in the spleen between neonates, 3-week and 3-month old animals are significantly increased in comparison to changes in MAIT expression between neonates, 3-week and 3-month old animals in PBMC and thymus (p=0.021).

Table 4-6: Results of statistical analyses of relative abundance of MAIT TCRα transcripts from various tissues (thymus, spleen and peripheral blood mononuclear cells (PBMC)) from calves in three different age groups (neonatal, 3-week old and 3-month old), and the interaction between tissue type and age group.

Effect / Interaction	p value
Tissue Type	< 0.001 ***
Age Group	0.0057 **
Tissue Type:Age Group	0.0216 *

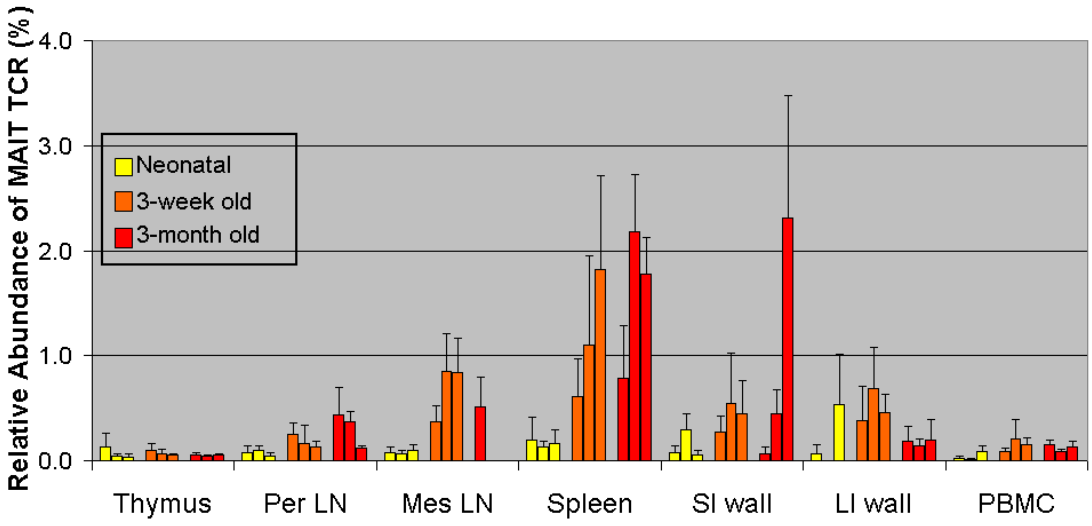


Figure 4-5: Relative abundance of bovine MAIT TCRα expression in thymus, peripheral lymph node (PerLN), mesenteric lymph node (MesLN), spleen, small intestinal wall (SIwall), and large intestinal wall (LIwall), and peripheral blood mononuclear cells (PBMC), from three animals within each of three age groups: neonatal, 3-week old, 3-month old. MAIT expression in the mesenteric lymph node was only tested in one of the three 3-month calves. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

4.3.3 Expression of ovine MAIT TCRα transcript is most abundant in prescapular lymph node, spleen and small intestine

The same qPCR assay, with specific ovine MAIT primers, was used to assess MAIT cell prevalence and distribution in sheep. Tissue samples from thymus, prescapular and mesenteric lymph nodes, spleen, and small and large intestinal walls, as well as circulating PBMC, were collected from two six-month old sheep. Samples from both animals were analysed twice and provided similar results. Overall levels of MAIT transcript were lower in comparison to cattle (0.1-0.5%), with highest expression of MAIT transcript (0.3-0.5%) in prescapular lymph node, spleen and ileum. Levels in thymus, colon and PBMC were all <0.1% (see figure 4-6).

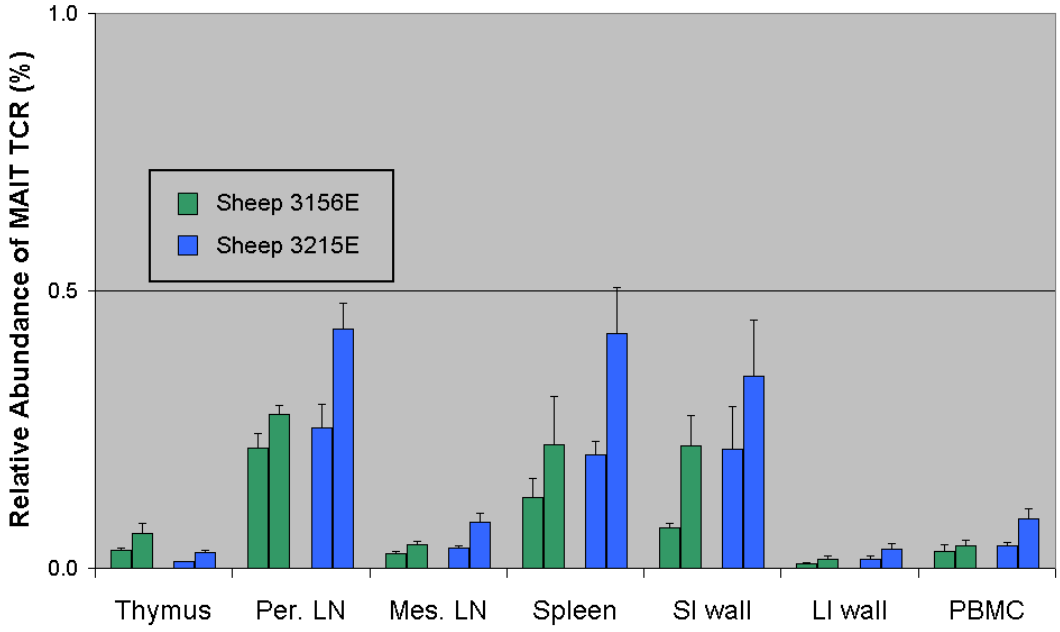


Figure 4-6: Relative abundance of ovine MAIT TCRα expression in thymus, peripheral lymph node (PerLN), mesenteric lymph node (MesLN), spleen, small intestinal wall (SIwall), and large intestinal wall (LIwall), as well as peripheral blood mononuclear cells (PBMC), from two six-month old sheep. Samples from both animals were analysed twice with similar results. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

4.3.4 $CD4^-$ $CD8^-$ $\gamma\delta^-$ T-cell populations were detected in PBMC from all age groups of cattle

On the basis of qPCR evidence that bovine MAIT cells increase in number within the first 3-weeks of life, PBMC from neonatal and 3-week old ("juvenile") calves were analysed by indirect immunofluorescence staining and flow cytometry to assess if the number of $CD4^-$ $CD8^-$ αβ T cells, as well as other T cell subsets, significantly changed during the post-natal period. Three calves were sampled as neonates only, and three were sampled as neonates and at three weeks of age. Six adult (>2-years old) cattle were sampled for comparison. Staining with specific antibodies was used to identify CD3, CD4, CD8, γδ, and DN ($CD3^+$, $CD4^-$, $CD8^-$, $\gamma\delta^-$) αβ T-cell populations. Results are presented in table 4-3. The anti-TCR γδ monoclonal antibody was accidentally omitted from the assays of one of the neonatal calves and therefore data from this animal were not included in the analyses.

The percentages of DN αβ T-cells were slightly higher in the neonatal calves (mean 2.9%), compared to the 3-week old calves (mean 1.2%). However, the values varied between individual animals within the groups (0.3% to 4.5% for neonates and 0.6% to 1.8% for 3-week old calves). Moreover, higher levels of DN αβ T cells were found in the in adult cattle (mean 3.2%, range 2.0-4.2%) compared to both of the younger groups. High levels of γδ T-cells (mean 40%, range 26.0-61.8%) were detected in all of the neonatal animals and although they were slightly higher than the levels in 3-week old calves (mean 36%, range 23-43%) this was not statistically significant. The percentages of γδ T-cells in adult cattle (mean 27%, range 19%-39%) were lower than in the younger groups. Levels of T-cells ($CD3^+$) in PBMC from different age groups were fairly similar, ranging from 63.3% to 79.7% in all except for one neonatal animal which had 46.6% $CD3^+$ cells. Levels of $CD4^+$ and $CD8^+$ cells were more variable between individual animals in the neonatal group than in the other groups, but among the 3 calves sampled after birth and at 3 weeks, those showing low percentages of $CD4^+$ and $CD8^+$ cells at birth had increased percentages at 3-weeks of age.

Table 4-7: Percentages of T-cell subpopulations in PBMC T-cell populations of neonatal, juvenile (3-weeks old) and adult cattle determined by immunofluorescence staining and flow cytometry. PBMC were isolated from blood collected by jugular venepuncture and data are expressed as the percentages of total PBMC.

		T-cell subpopulation (MAb used for staining)				
Age	**Animal**	**$CD3^+$** (MM1A)	**$CD4^+$** (IL-A12)	**$CD8^+$** (IL-A105)	**$\gamma\delta^+$** (GB21A)	**$CD3^+$ $CD4^-$ $CD8^-$ $\gamma\delta^-$**
Neonatal	**2511**	**63.7%**	20.8%	6.7%	29.3%	4.5%
	0312	**63.3%**	12.4%	11.8%	33.8%	1.4%
	1776	**68.2%**	33.7%	10.3%	26.0%	3.0%
	1777	**46.6%**	2.8%	2.4%	49.9%	2.6%
	1803	**74.0%**	6.1%	7.7%	61.8%	0.3%
3wks	**1776**	**63.8%**	31.3%	14.2%	22.9%	1.8%
	1777	**76.6%**	19.1%	11.8%	42.9%	1.2%
	1803	**66.4%**	15.4%	8.7%	42.4%	0.6%
Adult	**002**	**70.8%**	17.5%	6.2%	39.1%	3.8%
	332	**66.2%**	23.3%	9.9%	27.2%	4.0%
	475	**79.7%**	33.3%	19.5%	28.7%	2.0%
	481	**65.9%**	19.2%	14.2%	25.3%	2.7%
	851	**79.5%**	24.4%	11.0%	21.9%	4.2%
	860	**71.4%**	19.4%	29.1%	18.9%	2.0%

4.3.5 MAIT TCRα expression is highest in the bovine $CD8^+$ T-cell subset

Results of PCR and sequence analyses of expressed TCRα chains in fractionated T-cell populations, presented in chapter 3, suggested that MAIT cells may not be enriched within the $CD4^-$ $CD8^-$ population of bovine T-cells, as they are in humans and mice (Tilloy et al. 1999). To investigate this further, an initial experiment was conducted to compare the frequency of MAIT TCR transcripts in whole PBMC with that in positive and negative fractions obtained following staining with a mixture of antibodies specific for CD4 and CD8 and γδ TCR ($CD4^+$+$CD8^+$+$\gamma\delta^+$ and $CD4^-$ $CD8^-$ $\gamma\delta^-$ fractions). A standard PCR for MAIT transcripts gave readily detectable product

of the expected size from the PBMC and $CD4^+$+$CD8^+$+$\gamma\delta^+$ populations, with the latter giving a heavier band, whereas the $CD4^-$ $CD8^-$ $\gamma\delta^-$ fraction yielded much less product. However, the latter would be expected to contain much fewer T cells than the former two populations (see figure 4-7).

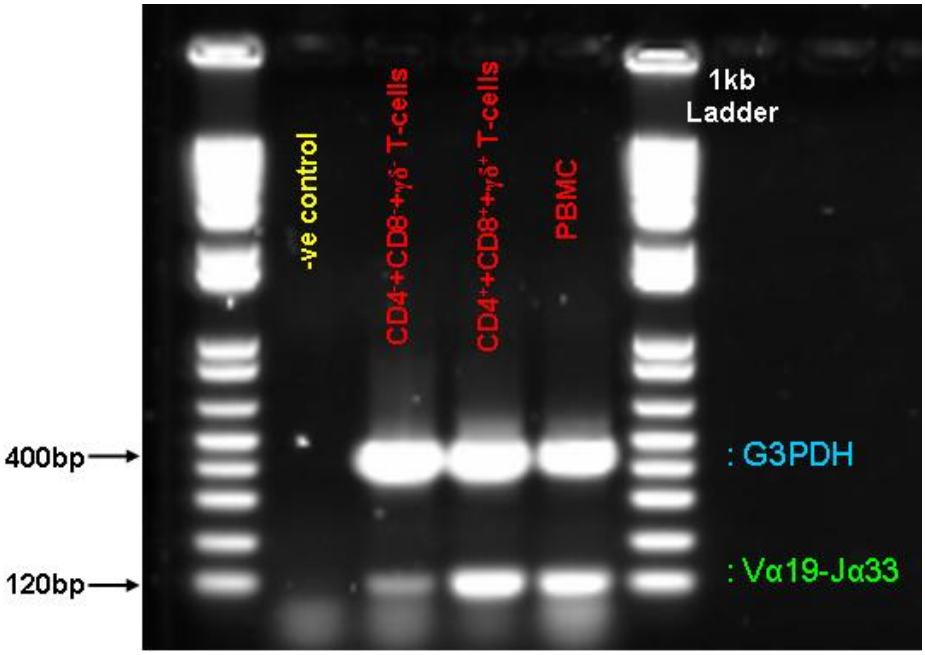


Figure 4-7: Agarose gel showing PCR amplification of MAIT transcript (Vα19-Jα33) in comparison to the house-keeping gene G3PDH in purified populations of bovine $CD4^-$ $CD8^-$ $\gamma\delta^-$ T-cells, $CD4^+$+$CD8^+$+$\gamma\delta^+$ T-cells, and unfractionated PBMC, as well as a negative control of sterile distilled water.

The same samples were analysed by qPCR to determine the abundance of MAIT TCRα in the αβ T cell populations. Based on their abundance relative to total TCRα, the PBMC, $CD4^+$+$CD8^+$+$\gamma\delta^+$ and $CD4^-$ $CD8^-$ $\gamma\delta^-$ populations were estimated to contain 0.64%, 0.82% and 0.29% MAIT transcripts respectively (figure 4-8). These results further indicate that ruminant MAIT cells are not enriched within the $CD4^-$ $CD8^-$ but rather that most of them are included in either the $CD4^+$ or $CD8^+$ populations. Although $\gamma\delta^+$ T-cells would have been included in the positive sorted population, the PCR would not be expected to amplify transcripts from these cells.

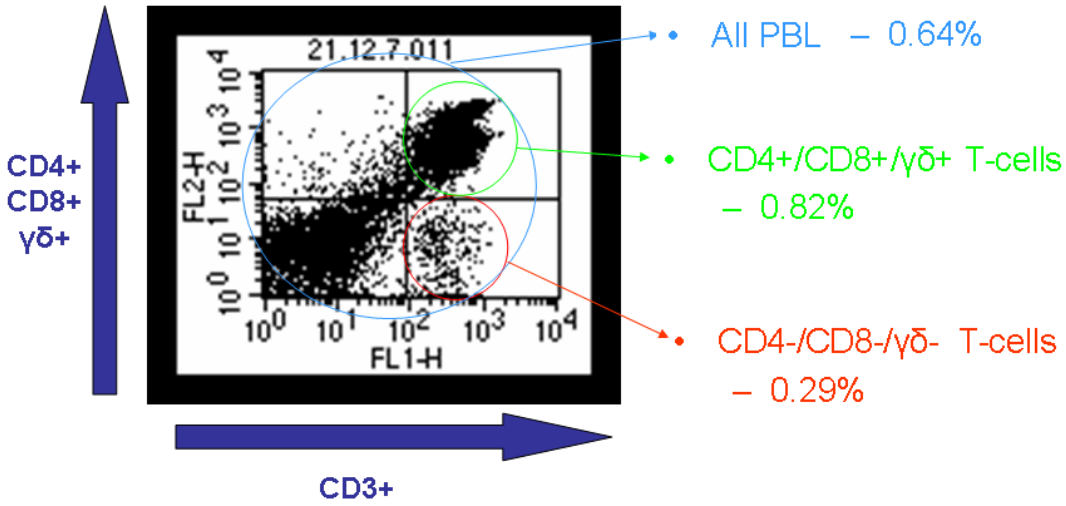


Figure 4-8: Relative abundance of MAIT TCRα expression in purified bovine T-cell subpopulations. T-cell populations were identified by primary antibody staining with MM1A (CD3), IL-A12 (CD4), IL-A105 (CD8), GB21A (γδ T-cells), and immunofluorescence staining using secondary isotype-specific antibodies, and flow cytometry analysis, and were purified by flow cytometry into a $CD4^+$+$CD8^+$+$γδ^+$ population (circled in green) and a $CD3^+CD4^-CD8^-γδ^-$ population (circled in red). Relative abundance of MAIT TCRα expression was assessed in each of these populations as well as unfractionated PBMC (circled in blue) by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool; results are shown in text on the right of the figure.

Samples of PBMC from a further 4 animals were tested to determine the extent of inter-animal variation in MAIT expression. Samples were tested twice with similar results each time. The average values for relative abundance of MAIT TCR were (animal)#109: 0.56%; #158: 0.10%; #1706: 0.22%; #1707: 0.29% (see figure 4-9). To determine whether ruminant MAIT cells are more prevalent within the $CD4^+$ or $CD8^+$ T-cell subpopulations, a series of cell sorts were undertaken on PBMC from one of these animals (109) to obtain $CD3^+$ and $CD3^-$ cells, $CD4^+$ and $CD4^-$ cells, and $CD8^+$ and $CD8^-$ cells.

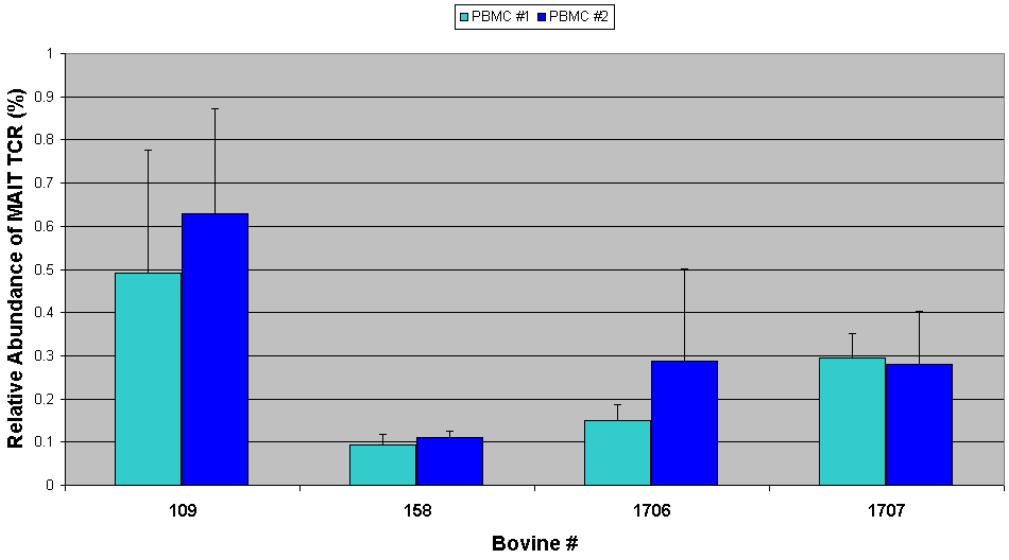


Figure 4-9: Relative abundance of MAIT TCRα expression in PBMC of 4 cattle. Samples were run twice for accuracy; the results from each experimental run are shown in the same colour. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

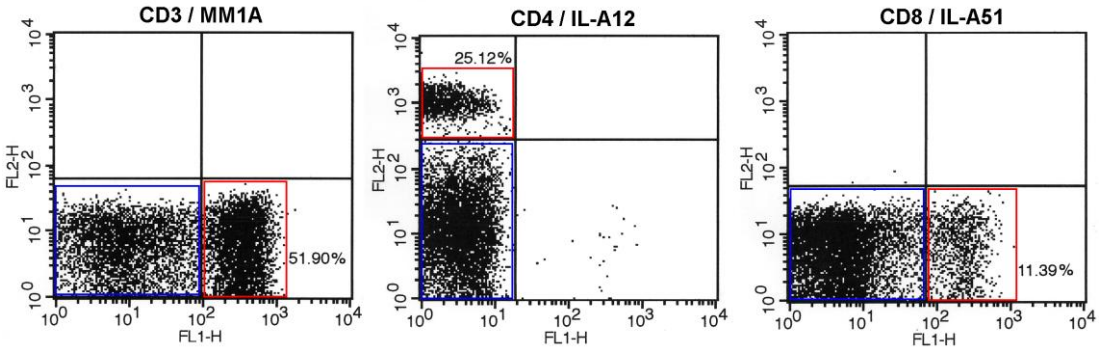


Figure 4-10: Purification of CD3+ and CD3-, CD4+ and CD4-, and CD8+ and CD8- populations from bovine PBMC by flow cytometry. Populations were identified by staining with primary antibodies MM1A (CD3), IL-A12 (CD4), IL-A51 (CD8) and immunofluorescence staining using secondary isotype-specific antibodies, and gated as positive (red) or negative (blue). Figures are given for the percentage of PBMC identified as staining positive for each antibody.

The cell sorts yielded between 8 x 10^5 and 4 x 10^6 cells for each fraction and re-analysis demonstrated that they were 98-99% pure. qPCR analysis of equal amounts of cDNA extracted from these fractions indicated that MAIT transcripts were enriched within the $CD3^+$ ($CD3^+$: 0.8%; $CD3^-$: 0.2%) and also within the $CD4^-$ ($CD4^+$: 0.0%; $CD4^-$: 2.3%), and $CD8^+$ ($CD8^+$: 2.5%; $CD8^-$: 0.1%) populations (see figure 4-11). These results indicate that bovine MAIT cells are predominantly found in the $CD3^+$ $CD4^-$ $CD8^+$ population. The detection of 0.2% MAIT expression in the $CD3^-$ population probably relates to $CD3^+$ cells which have low mean fluorescence intensity (MFI) of MM1A staining and have been accidentally purified as $CD3^-$ due to the position of flow cytometry sorting gates at a relatively high MFI.

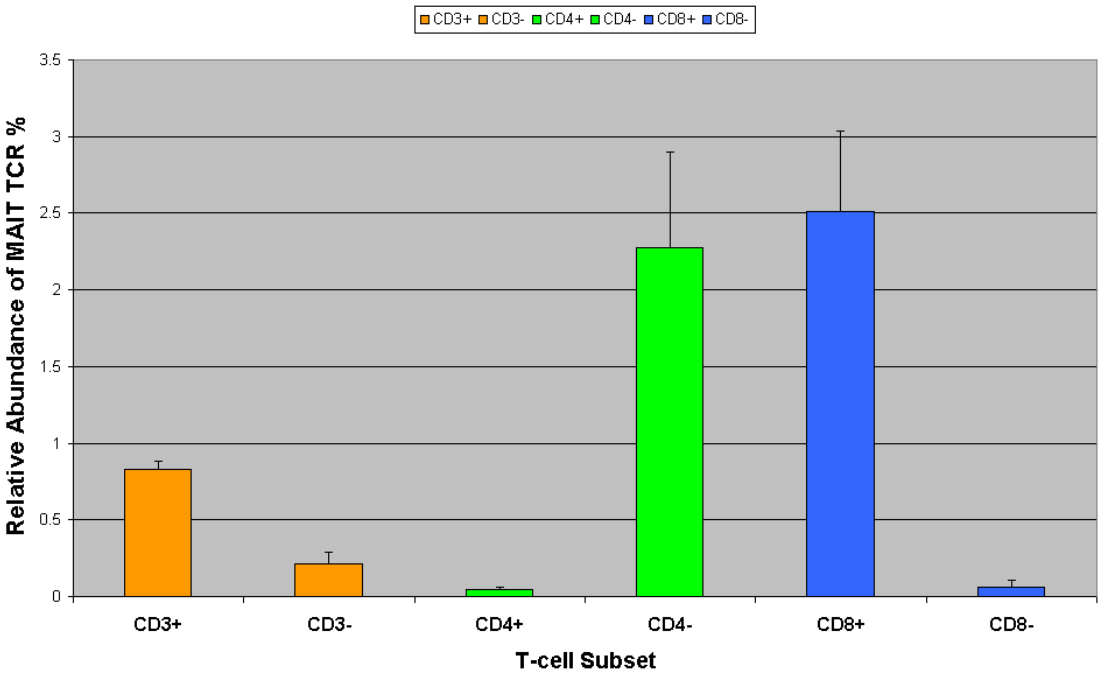


Figure 4-11: Relative abundance of MAIT TCRα expression in bovine T-cell subsets purified as positive and negative for each phenotype by flow cytometry and staining with primary antibodies MM1A (CD3), IL-A12 (CD4), IL-A51 (CD8) and immunofluorescent isotype-specific secondary antibodies, as shown in figure 4-10. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. Error bars represent the standard error of each result.

4.3.6 Analysis of MAIT TCRVα expression in sheep infected with *Telodorsagia circumcincta*

To test the applicability of the MAIT qPCR to studies of responses to infection, the assay was used to assess MAIT transcript abundance in the efferent gastric lymph of sheep undergoing experimental infection with *Telodorsagia circumcincta*. This is a common gastrointestinal parasite which colonises the abomasal mucosa, causing chronic ill-thrift and reduced productivity. Two groups of previously worm-free animals were studied, one subjected to a primary infection and the other to a secondary challenge following immunisation by previous trickle infection and anthelmintic treatment. Samples of cDNA prepared from lymph samples collected on the day before challenge, and on days 1, 3, 7, 14, 21 in naïve animals, and days 1, 3, 5, 7, 9 in previously-infected animals were analysed for relative abundance of MAIT TCRα by qPCR. The results obtained were compared to the abomasal worm counts obtained following slaughter of the animals on day 21 or day 9, respectively.

Results from analyses are presented in figures 4-12 and 4-13, and combined in figure 4-14. As observed in PBMC and tissues from healthy sheep, the estimated levels of MAIT transcript in gastric lymph were low throughout the experiments (<0.2%). In general, animals undergoing primary infection showed a small increase in MAIT abundance after challenge, which was still increased at day 21; this was greatest in animals B103 and 955 (0.05% on day -1 to 0.15% on day 21, and 0.03% on day -1 to 0.13% on day 21). Previously-infected animals had relatively low abundance of MAIT transcript before challenge (<0.04%) and also at day 9 after challenge (<0.04%), but 4 of the 5 animals showed a small increase at day 3 after challenge. Overall, however, although MAIT transcripts exhibited these small changes in abundance, these were inconsistent and not statistically significant. Terminal abomasal worm burdens were variable and did not significantly correlate with changes in MAIT abundance within either group.

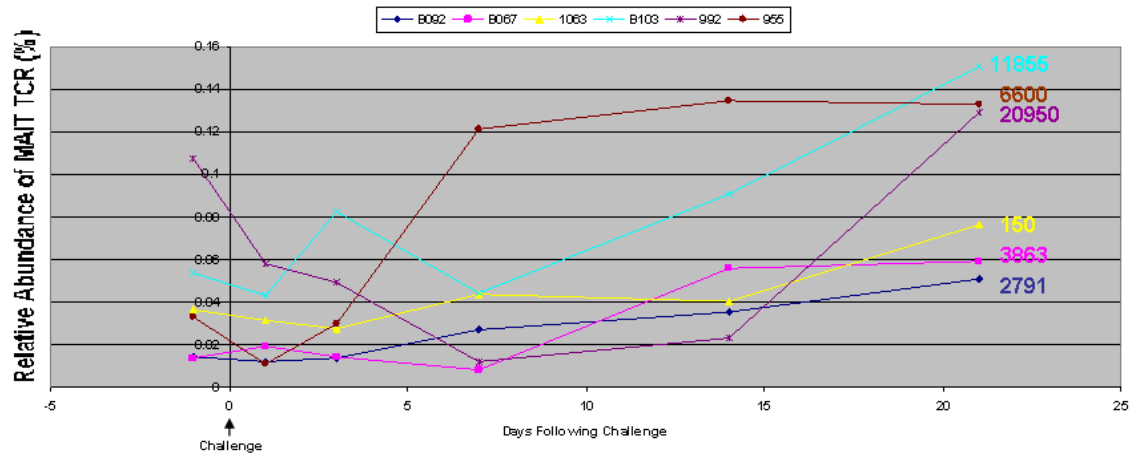


Figure 4-12: Relative abundance of MAIT TCRα expression in efferent gastric lymph of sheep undergoing primary infection with the gut nematode, *Telodorsagia circumcincta*, before and following challenge on day 0. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. The standard error of each result is not shown.

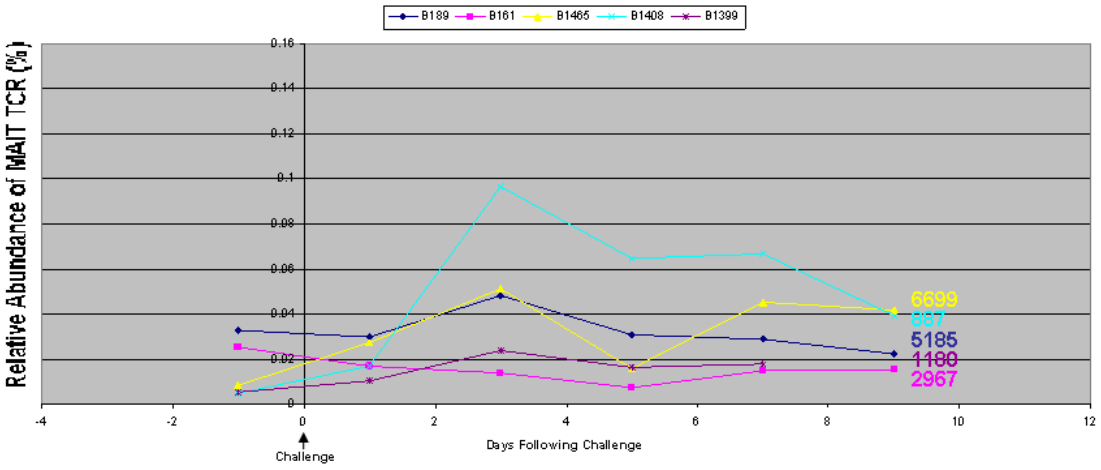


Figure 4-13: Relative abundance of MAIT TCRα expression in efferent gastric lymph of sheep undergoing secondary infection with *T. circumcincta*; these had been previously-infected with the same parasite but treated with anthelmintic before repeated challenge. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. The standard error of each result is not shown.

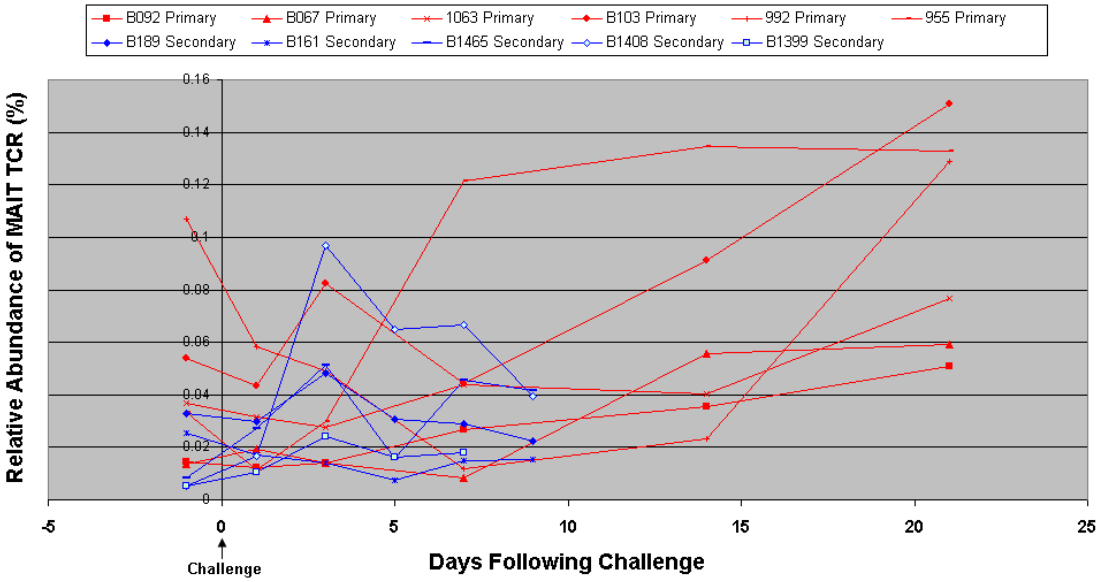


Figure 4-14: Relative abundance of MAIT TCRα expression in efferent gastric lymph of sheep undergoing primary (red) or secondary (blue) infection with the gut nematode, *Telodorsagia circumcincta*, before and following challenge on day 0. Relative abundance of MAIT TCRα was calculated by using quantitative PCR to detect MAIT TCRα expression in comparison to the total TCRα pool. The standard error of each result is not shown.

4.4 Discussion

A real-time quantitative PCR was designed to examine relative abundance of MAIT cells, using primers specific for the Vα and CDR3 of the MAIT TCRα chain, together with a second reaction to measure total TCRα transcripts, using constant region-specific primers. This provided estimates of MAIT TCRα expression in about 0.1-0.6% of T-cells in PBMC of young adult cattle, and in 0.1-0.2% of T-cells in PBMC of 6-month old sheep. These figures might represent an underestimate, as a proportion of T-cells in mice and humans have been shown to express two re-arranged TCRα transcripts (Heath and Miller 1993; Padovan et al. 1993). Nevertheless, since preliminary analysis of the bovine TCRα locus has identified >300 Vα genes (Reinink and Van Rhijn 2009), these estimates of MAIT TCRα abundance are still far in excess of that expected for sequences generated by random rearrangement in the absence of further expansion.

Previous studies have concluded that MAIT cells undergo development in the thymus, as they were absent in athymic mice. However, MAIT expression was not detected in human cord blood lymphocytes, nor in the thymus or spleen of neonatal mice, using RT-PCR (Tilloy et al. 1999). The inability to detect these cells in the thymus suggested that they may be produced in very small numbers and subsequently undergo peripheral expansion. This would be consistent with the finding that MAIT cells have an oligoclonal TCR Vβ repertoire (Treiner et al. 2005). In support of this suggestion, recent studies using a monoclonal antibody specific for human MAIT Vα7.2 demonstrated the presence of small numbers of MAIT cells with a naïve T cell phenotype within the thymus, and larger numbers of circulating MAIT cells with a memory phenotype in peripheral lymphoid tissues (Martin et al. 2009). Studies in B-cell deficient and germ-free mice have shown that peripheral expansion of the MAIT cell population (but not their production in the thymus) is dependent on B-cells, as well as an intact gut flora (Treiner et al. 2003).

To determine whether the numbers of MAIT cells in the peripheral T-cell pool also increase following birth in cattle, the qPCR for relative abundance of MAIT TCRα

was applied to samples from neonatal, 3-week old, and 3-month old calves. MAIT TCRα expression was uniformly low in all tissues in neonatal animals, but highly statistically-significantly increased in 3-week old and 3-month old calves compared to the neonates. Further analyses showed that the increase in MAIT expression with age was statistically significant in a number of secondary lymphoid tissues but was particularly pronounced within the spleen, and was also significantly increased compared to the thymus within the same animal. The paucity of MAIT transcripts in neonates, and subsequent rapid rise in MAIT transcript numbers with age, supports the model that this cell population undergoes expansion following birth. Further, the consistently low levels of MAIT TCRα expression in the thymus suggest that this expansion occurs extra-thymically. Given the evidence in mice that MAIT cell expansion is dependent on the gut flora (Treiner et al. 2003), expansion of these cells in cattle is likely to occur in the immediate post-natal period, since previous work has demonstrated that maximal numbers of total faecal flora are reached within 24 hours of birth (Gay 1965). Levels of MAIT transcripts showed no significant increase between 3-week and 3-month old calves, and were consistent with older animals, suggesting that the initial population expansion was complete by 3-weeks of age. A rapid establishment of the MAIT cell population would be appropriate if these cells do possess innate function, as the innate immune system is particularly important during this post-natal period. If, as in mice, MAIT population expansion in cattle is dependent on B cells, this cell population is present in neonatal calves, although in smaller numbers than in older animals (Kampen et al. 2006). Rapid activation and expansion of the B cell population in response to microbial challenge is also expected to occur following birth.

Analyses of MAIT cell tissue distribution in cattle and sheep demonstrated predilections for similar tissues, particularly the spleen and some gut-associated tissues. This is the first evidence that MAIT cells represent a prominent T-cell population in the spleen. A previous study had failed to demonstrate MAIT expression within the spleen of neonatal mice (Tilloy et al. 1999); however, in this regard, the B-cells in the spleen of clean, healthy mice contain much fewer activated B-cells and there are very few germinal centres compared to other species living in

their natural environment. This finding raises important questions concerning the role of the gut as a site of expansion of these cells. The finding that expansion of the MAIT cell population in mice is dependent on a gut flora (Treiner et al. 2003) has prompted the suggestion that a microbial-derived MR1-associated ligand may be involved. Alternatively, the role of the gut flora may be to activate and expand B-cell populations, which through migration of the B-cell progeny, is likely to have an impact on B-cell populations in systemic lymphoid tissues. In this model, expansion of MAIT cells might occur in response to endogenous ligands expressed by activated B-cells. In this regard, the spleen of ruminants is relatively rich in B cells and in active germinal centres.

qPCR assays showed MAIT cells to represent 0.1-0.2% of all human peripheral T-cells. These are approximated to be 5-10 times more abundant in humans than mice, since they compose 15% of human DN T-cells in comparison to 2% of murine DN T-cells (Tilloy et al. 1999; Treiner et al. 2005). However, a definitive quantification of murine MAIT cells in PBMC has been difficult to perform due to their paucity, and this result is only approximate since the size of the DN population and the proportion of MAIT cells within the DN population may differ between the species. Certainly, recent studies demonstrated that circulating MAIT cells in mice have a naïve phenotype, and it has been hypothesised that MAIT do not undergo peripheral expansion in this species (Martin et al. 2009). This study used antibody against the human Vα7.2 gene to identify MAIT cells and suggested that these may represent up to 4% of human peripheral T-cells. The estimates of frequency of MAIT cells in cattle, based on quantifying expression of the MAIT TCRα chain, ranged from 0.1% to 0.6% of αβ T cells in PBMC in different animals, and hence were similar to frequencies determined by the same methods in humans. Interestingly, these levels were roughly five times higher than those obtained in sheep where MAIT represented only 0.01-0.08% of PBMC (i.e. more similar to the levels in mice). It has been proposed that MAIT cells may perform a similar functional role to iNKT cells which are restricted by the class I-like molecule CD1d (Godfrey et al. 2000). While humans have higher numbers of MAIT cells than mice, the converse is the case for iNKT cells, with mice having 10-fold more of these cells (Treiner and Lantz

2006). In this regard, sheep have been shown to possess a functional CD1d gene (Rhind et al. 1999), whereas genome searches have failed to identify a functional CD1d gene in cattle and cattle T-cells do not stain with CD1d-ligand tetramers (Van Rhijn et al. 2006). Hence it is possible that the lower levels of MAIT cells in sheep compared to cattle may be related to differences between the species in the relative functional roles of these invariant T-cell populations. Further comparative knowledge of the characteristics and functions of these cell types would be required to address this hypothesis.

MAIT cells were originally identified as being enriched in the small $CD4^-$ $CD8^-$ T cell population in mice and humans and preliminary results indicated that this was also the case in cattle (Tilloy et al. 1999). In humans, there was evidence that they were also present in the $CD8\alpha\alpha^+$ population, and conversely some enrichment was seen in the murine $CD4^+$ population (Tilloy et al. 1999). However, results from sequencing of TCRα cDNAs isolated from T cell subsets in the present study (reported in chapter 3) suggested that this might not be the case in ruminants. Thus, no MAIT TCRα sequences were found in 90 cDNAs examined from a $CD4^-$ $CD8^-$ $\gamma\delta^-$ purified population whereas 2 MAIT sequences were identified in a similar sample of cDNAs from unfractionated PBMC. To provide more quantitative data on the distribution of MAIT cells in bovine T-cell subsets, the qPCR for relative abundance of MAIT TCRα was applied to samples of cells purified based on expression of T-cell surface markers. In an initial experiment, results again demonstrated that MAIT expression was higher in cells sorted as positive following staining with a mixture of antibodies specific for CD4, CD8 and γδ TCR, than in the corresponding negative population. Two-colour staining of PBMC from a number of cattle demonstrated that $CD4^-$ $CD8^-$ T cells usually represent <4% of PBMC, T-cells representing >50% of PBMC. This finding, together with the MAIT qPCR data, indicates that a large majority of the MAIT cells are within the $CD4^+$ and/or $CD8^+$ populations. Application of the MAIT qPCR in a second experiment confirmed the low representation of MAIT sequences in the $CD4^-$ $CD8^-$ γδ TCR^- population and demonstrated that they are enriched in purified $CD8^+$ cells compared to $CD8^-$ cells and in $CD4^-$ cells compared to $CD4^+$ cells. These results indicate that bovine MAIT

cells are found predominantly within the $CD8^+$ T-cell population, thus differing from humans and mice in which they are most enriched in the $CD4^-$ $CD8^-$ T-cell population.

These results suggest that bovine MAIT – MR1 interaction and activation may be more reliant on the CD8 molecule as a co-receptor than in humans or mice. However, recently published studies have demonstrated the presence of human MAIT cells within the CD8αβ population of T-cells as well as the DN and CD8αα T-cell subsets. These authors suggested that MAIT cells bearing a CD8αβ molecule had been missed during previous flow cytometry sorting due to low levels of staining for CD8β (Martin et al. 2009). IL-A51 and IL-A105 antibodies used to detect bovine CD8 molecules are specific for the CD8α chain and therefore detect both CD8αα and CD8αβ molecules (MacHugh and Sopp 1991). A monoclonal antibody specific for the bovine CD8β chain, CC-58, has been described and used in two-colour staining to identify T cells expressing CD8αα from those expressing CD8αβ (MacHugh and Sopp 1991). Further studies using monoclonal antibodies specific for the different forms of CD8 chain are required to determine whether bovine MAIT cells are enriched in CD8αα or CD8αβ populations.

The qPCR assay designed to examine relative abundance of MAIT cells proved a useful tool for analysis of MAIT TCRα expression in relation to the total αβ T-cell pool. Results were consistent between triplicates and repeated experiments, with relatively low levels of standard error. Detection of MAIT and TCR transcripts was highly sensitive, and allowed detection down to, and including, $1x10^{-10}$ng/µl. However, the assay also had its limitations. In particular, because MAIT cells are usually present in small numbers, and since MAIT transcripts may not correlate exactly with cell numbers, the qPCR gives only an approximate estimate of the numbers of cells present. A number of samples were also near the lower threshold for detection of MAIT transcript, and for one sample (LI wall of calf 0312) no MAIT amplification was achieved. Also, since the assay is based on analysis of cDNA, then increased activation status of αβ T cell populations other than MAIT cells would proportionately increase overall TCRα transcript relative to MAIT transcript, and

result in an underestimation of MAIT cell number. In conclusion, although this qPCR assay is useful to compare the representation of MAIT cells in different populations, it has limitations in providing precise data on the numbers of these cells. More detailed studies would be possible by producing a monoclonal antibody specific for the MAIT TCRα. This would allow accurate identification of MAIT, and purification of these cells for functional studies. However, since the studies described in chapter 3 showed that use of the germline Vα19 segment is not exclusive to the MAIT TCR, an antibody specific for the MAIT CDR3 loop would be required to avoid cross-reactivity with other T cells expressing a TCRα chain bearing Vα19.

Since MAIT cells have been previously demonstrated to be increased in abundance in the intestinal mucosa of humans and mice, the qPCR assay for MAIT TCRα was applied to examine changes in MAIT abundance in sheep infected with the gastro-intestinal parasite *T. circumcincta*. Samples of efferent gastric lymph draining from the abomasum of experimentally-infected sheep were available for the study. This allowed assessment of the suitability of this assay as a tool to investigate changes in MAIT cell abundance during responses to infection. No marked change in relative abundance of MAIT cells was detected following either primary or secondary infection. This may relate to retention of MAIT cells in the lymph node, or in the gut mucosa, in which case changes would not be detected in efferent lymph. Alternatively, an increase in abundance of MAIT cells may not have been detectable because of similar increases in numbers of activated cells within other T-cell subsets.

In conclusion, the work in this chapter shows that bovine MAIT cells are present at a similar frequency to that observed for human MAIT cells, but in contrast to humans are most abundant within the $CD8^+$ T cell subset rather than the $CD4^-$ $CD8^-$ population. Analyses of tissue distribution and changes with age suggest that bovine MAIT also undergo extra-thymic peripheral clonal expansion. Of particular interest was the observation that MAIT cells were most abundant in the spleen. The relatively high frequency of MAIT cells in peripheral lymphoid tissues, coupled with the high level conservation of MAIT TCRα chain CDR sequences, support the idea

that these cells may participate in innate immune responses by recognition of a similarly conserved ligand.

5 Molecular Cloning and Characteristics of Ruminant MR1

5.1 Introduction

The highly polymorphic classical class I genes involved in presentation of antigenic peptides to T-cells are located within the MHC locus. The human MHC region is sited on chromosome 6 (The M. H. C. sequencing consortium 1999), and contains classical and non-classical MHC class I-like genes; the latter include a number of non-polymorphic class I genes, e.g. HLA-E in man, which presents ligands to NK cells and NK-CTL. Additional class I-like genes are found at other locations in the genome – these include the CD1 genes and MR1, which in humans are found on chromosome 1q23 and 1q25 respectively, within an MHC-paralogous region (Kasahara 1999). Studies with mice in which the MR1 gene has been deleted indicate that development of MAIT cells is dependent on MR1, and in vitro studies suggest that MAIT cells respond to an unknown ligand presented by MR1 (Treiner et al. 2003).

Orthologous MR1 genes have been detected in a number of different mammalian species and demonstrate a striking level of cross-species sequence conservation at both nucleotide and amino acid level. The extracellular domains of human and murine MR1 share up to 90% identity, particularly the α1 and α2 domains which form the ligand binding site (Yamaguchi et al. 1997). This is the highest level of inter-species identity of all known class I-related molecules and implies an important immunological function which evolved prior to species divergence (Yamaguchi et al. 1997; Treiner et al. 2003). The MR1 gene is notably monomorphic (Parra-Cuadrado et al. 2000) and has an unusually large first intron in comparison to other class I genes (Yamaguchi et al. 1998); it also demonstrates 'type I splicing', where each domain starts with a composite codon whose first nucleotide is contributed by the previous exon (Riegert et al. 1998). A number of alternatively-spliced transcripts of MR1 have been reported in humans and mice, in addition to full MR1 sequence; the majority of these involve sequence disruption in the α1 (murine) or α3 (human) regions (Riegert et al. 1998).

The predicted structure of MR1 suggests that it shares structural features with classical MHC class I molecules, notably the characteristic MHC class I protein fold of the heavy chain comprising 3 extracellular domains, α1, α2 and α3, and an ability to associate with β2-microglobulin (Yamaguchi et al. 1997; Yamaguchi and Hashimoto 2002). This also includes a typical groove, formed by the α1 and α2 domains, that is predicted to accommodate a ligand; however, the nature of this ligand has yet to be discovered. Molecular modelling suggests that the binding groove of MR1 is hydrophilic, but comparative analysis to classical class I molecules shows differences in MR1 within five of the eight class I residues defined to contact peptide termini (Madden 1995), suggesting that MR1 is unlikely to bind to a peptide ligand (or if so, it does so in a novel way) (Riegert et al. 1998). Additionally, MAIT cells have been shown to be activated by synthetic glycolipids (Okamoto et al. 2005), suggesting that MR1 may present a lipid/glycolipid ligand to MAIT, in a manner similar to that presented by CD1d to iNKT cells.

Analyses of MR1 expression in human and murine tissues by RT-PCR or Northern blotting have demonstrated that it is transcribed in a wide range of tissues and in cell-lines of different lineages (Hashimoto et al. 1995; Riegert et al. 1998). However, cell surface expression of MR1 has yet to be demonstrated *in vivo*, and following transfection into cell lines *in vitro*, MR1 is expressed internally but usually only at low levels on the cell surface (Huang et al. 2005; Treiner et al. 2005). Studies of MR1-transfected cell lines demonstrated that most MR1 protein was retained intra-cellularly and analyses of subcellular localisation also indicated that MR1 is transported through the endosomal compartment, leading to the suggestion that it acquires its ligand within the endosomes (Huang et al. 2008).

The work described in this chapter aimed to identify and characterise MR1 genes in cattle and sheep to provide further data for cross-species comparisons and for subsequent studies of MR1 expression.

5.2 Materials and Methods

5.2.1 Experimental animals

Samples of peripheral blood were obtained from healthy adult Holstein-Friesian cattle and Blackface-cross sheep as described in section 2.1. Peripheral blood mononuclear cells (PBMC) used for total RNA extraction and cDNA synthesis (see sections 2.4.1.1 and 2.4.2.1) were prepared as in section 2.3.1.1.

5.2.2 Online resources and sequence analysis

Online sequence databases from the National Center for Biotechnology Information (NCBI), including Entrez Nucleotide, and Ensembl were used to obtain original sequence data and search for published gene data. Sequence analysis was performed using DNASIS software (Hitachi Software Engineering Company, Ltd.) (see section 2.5).

5.2.3 PCR primers and reaction conditions

The primer sequences used for PCR amplification of bovine and sheep MR1 are shown in table 5-1. PCR were performed using reaction protocols and reagents as described in section 2.4.2.2. Optimal cycling conditions were assessed for each individual reaction; based on results, a standard reaction was then used, consisting of 10 minutes at 95°C, 35 cycles of (1 minute at 95°C, 1 minute at specific annealing temperature, 1 minute at 72°C), and a final extension period of 10 minutes at 72°C for all reactions except that using primers “ovMR1f” and “ovMR1r”, where a total of 40 cycles of (1 minute at 95°C, 1 minute at 54°C, 1 minute at 72°C) were performed. Optimal annealing temperatures for each primer combination are listed in table 5-1. PCR products were analysed on 1.5% agarose gels, as described in section 2.4.3.3.

Table 5-1: Primer pairs and optimal annealing temperatures used to amplify partial or full-length coding sequences of bovine and ovine MR1.

Target Sequence	Forward 5' sequence 3'	Reverse 5' sequence 3'	Optimal Tm
Full-length coding sequence bovine MR1 (original)	bovineMR1f CAGAGGACGAGACAATCGAAG	bovineMR1r AGATGGCTGCTTTCCAACTG	58°C
Full-length coding sequence bovine MR1 (Van Rhijn*)	iboMR1f GGGCTGATGATGCTCCTATTGCC	iboMR1r GAACACTGATGGGAGATGGCTGC	60°C
Ovine MR1 α2 region	boovMR1f2 CAGAGAATGATTGGCTGTGA	boov(n)MR1r TGGAAGAAGAATGCATTGC-	52°C
5' partial ovine MR1	ovMR1f AGGATTCTAGAGGGGCTGAT	ovMR1r CTTTACTGCAGAGCTCATGG	54°C
PCR to assess splice-variant structure:			
(i) 5' Untranslated region to α2	boMR1a2-f1 AAGCAGCTGCAGCATCACTA boMR1a2-f2 CATCACGGGATTTCTCCAAT	boovMR1r CCTCCTCTGTCCCATATG boovMR1r -	54°C 54°C
(ii) α2 to 3' Untranslated region	boovMR1f CGGAGAGAGCAGTACAAG boovMR1f -	boMR1a2-r1 GACATACTCAGGCGGGTATG boMR1a2-r2 TGGAGTATGGGAAAGATGCC	54°C 54°C

* **These primers were made available by Dr Ildiko van Rhijn, University of Utrecht, during the course of this project.**

5.2.4 Analyses of PCR products

DNA from PCR products was either purified directly (see section 2.4.3.5) or following extraction from bands in agarose gel (2.4.3.4). The DNA was cloned into pGEM-T Easy vector and used to transform Promega JM109 high efficiency competent cells (see section 2.4.4). Transformed *E. coli* colonies were selected on the basis of antibiotic resistance and absence of beta galactosidase, and tested by PCR using either pGEM-T Easy primers, T7 and SP6, or primers specific to the insert. Colonies of interest were grown overnight in LB medium and plasmid DNA extracted (see section 2.4.1.3). Sequencing of plasmid DNA from single clones was performed by DBS Genomics, Durham University, UK (section 2.4.3.6).

5.3 Results

5.3.1 Amplification of multiple bovine MR1 transcripts

Genomic and mRNA sequences of human and murine MR1 were used to search the NCBI online bovine gene databases, using the basic local alignment search tool, BLAST. A bovine expressed sequence tag (EST) which matched these was identified via NCBI Entrez Nucleotide; this showed >80% nucleotide identity to the 5' end of human and murine MR1 sequence, and was taken to represent partial bovine MR1 sequence. Cross-species comparison showed this sequence to represent leader, α1, and α2 exons, as well as a short part of the α3 exon (see figure 5-1 for representative diagram). This sequence was compared to the NCBI Nucleotide bovine and ovine gene databases using the website's Discontiguous MegaBLAST tool. No matches were detected in either *Ovis aries* (EST) or (other) databases. However, a large number of positive matches were found in *Bos taurus* databases – 6 expressed sequence tags (EST), 22 whole genome shotgun (WGS) sequences, and 7 'other' genes. Of these, six WGS, four EST, and 3 'other' gene matches bore close identity to bMR1. Eight of these closely matched the partial 5' bMR1 sequence and included additional 5' untranslated sequence, while four also contained sequence with close identity to the hMR1 and mMR1 α3 and trans-membrane (TM) exon sequences. One of these sequences, 963575, also included a segment corresponding to hMR1 and mMR1 cytoplasmic domain (73% and 58%, similarity respectively) and included a stop codon.

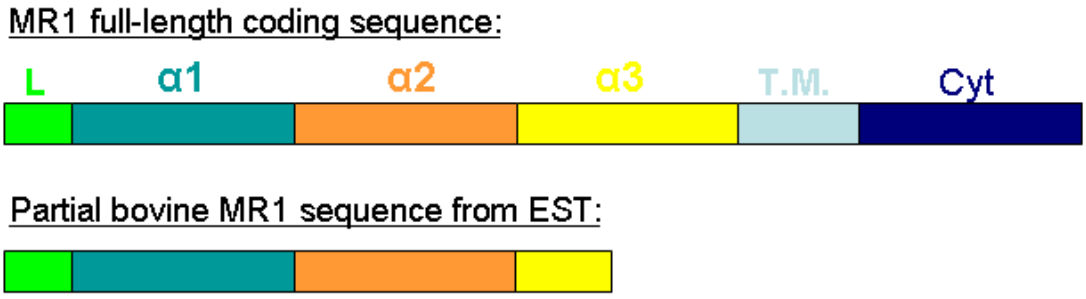


Figure 5-1: Schematic representation of MR1 sequence organisation. A bovine expressed sequence tag which showed a high level of similarity to human and murine MR1 was composed of leader, α1, α2, and a small section of α3 exons. L: Leader; TM: Transmembrane exon; Cyt: Cytoplasmic 'tail'.

Sequence from *Bos taurus* gene matches to MR1 (963575 and KN510_9508b20.p1km13F, NCBI accession numbers: CK774785 and FE016126) were combined with known bMR1 data to assemble a putative full-length bovine MR1 sequence. This sequence shared a high level of nucleotide identity to human and murine MR1 (76-88% in extracellular regions), and blast-searches within a preliminary assembly of the bovine genome database (Btau_3.1) using NCBI BLAST revealed an exact match within NW_001501825.1/BtUn_WGA3152_3. However, no match was made to the cytoplasmic domain, indicating that this part of the gene was missing from the genome assembly. (Full analysis of MR1 location within the *Bos taurus* genome is provided in section 5.3.3)

This putative bovine MR1 gene sequence was used to identify sites within the untranslated regions for location of PCR primers to amplify the full-length MR1 coding region. Primers "bovineMR1f" and "bovineMR1r", designed within the (5') and (3') untranslated regions, respectively, were tested on cDNA from bovine and ovine PBMC. Products were obtained in PCRs utilising a range of annealing temperatures. Several bands were visible in agarose gels. Bovine cDNA yielded products of 1100bp and 900bp, the former corresponding to the predicted full-length MR1 sequence, as well as a less abundant product of 650bp and some smaller weaker

products. The ovine PCR produced a range of smaller, less well defined bands, at 850bp, 725bp, 550bp, 320bp and 280bp.

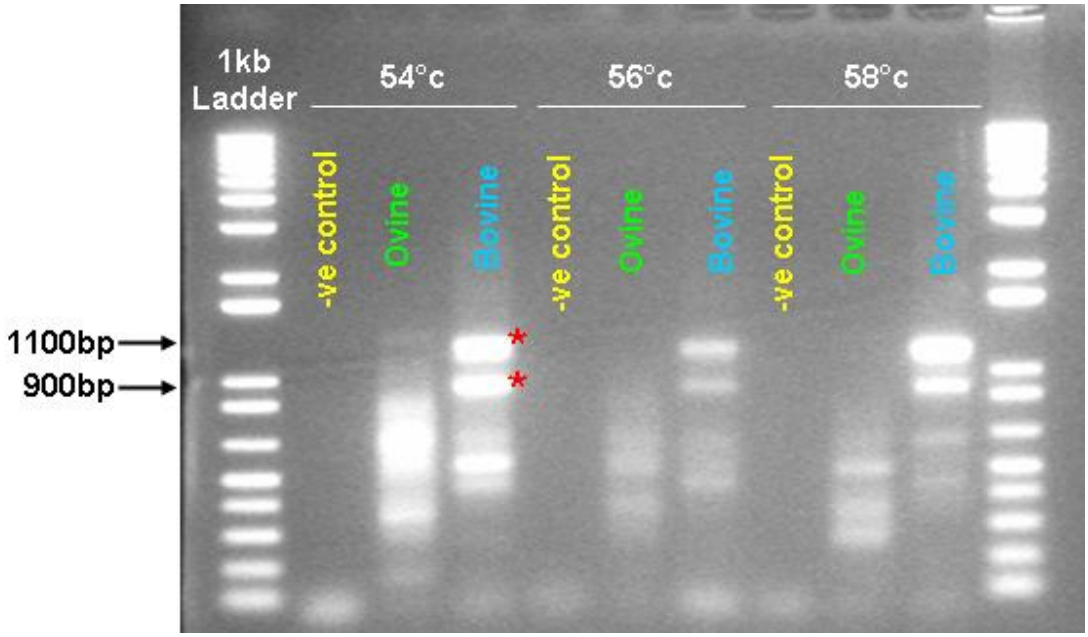


Figure 5-2: Agarose gel showing consistent amplification of multiple PCR products from bovine and ovine PBMC cDNA using primers boovMR1f/r at a range of annealing temperatures. Sterile distilled water was used as a PCR negative control. Bands excised from the gel for sequencing are shown by red asterisks.

Previous publications (Riegert et al. 1998) had identified a number of alternatively-spliced forms of human and murine MR1 of varying length. To determine whether the multiple PCR bands obtained from cattle and sheep also represented alternatively spliced transcripts, DNA was extracted from each distinct band (bovine: 1100bp, 900bp, 650bp; ovine: 850bp, 725bp, 550bp), and cloned into pGEM-T Easy vector. Sequencing of cloned inserts confirmed that the 1100bp bovine product shared 100% homology with the putative sequence assembled from EST and genome partial sequences, thus confirming the full-length bovine MR1 sequence. Sequences obtained from clones containing the 900bp and 650bp inserts were found to represent

different splice variants of bovine MR1. These sequences were missing either the α1 (900bp), or the α1 and α3 (650bp) exons, and are discussed further in section 5.3.5. Bovine MR1 sequence data is provided in Appendix D. Sequence data obtained from clones containing ovine products were of poor quality, and did not demonstrate any similarity to bovine MR1 sequence, suggesting that these did not represent ovine MR1.

During these experiments, communication with Dr. Ildiko Van Rhijn, University Utrecht, Netherlands, revealed that she was investigating MR1 and had designed a different set of primers, iboMR1f and iboMR1r, for bovine MR1 based over the 5' start codon and 3' stop codon, respectively. These primers were compared to the original bovine MR1 primer pair MR1f/r on cDNA extracted from PBMC of three different cattle. Results demonstrated marked differences between animals in both the intensity and number of products obtained with the two sets of primers – bovineMR1f/r produced multiple bands including the predicted full-length sequence from PBMC of animals 216 and 481, but no product from animal 002; in contrast, iboMR1f/r produced two strong products from animal 002, representing full-length MR1 and the splice variant missing the α1 exon, but weaker products from animals 216 and 481 (see figure 5-3). This experiment was repeated with similar results. Sequencing of the prominent bands (1100bp and 900bp) confirmed that the products obtained with both primer pairs represented MR1 and had identical nucleotide sequences. These results suggest that there may be sequence polymorphism between animals in one, or both, of the primer sites in the untranslated regions used by the bovine MR1f/r PCR (bovineMR1f, for example, is >50bp from the initiation codon). However, no differences were detected between sequences from different animal PCRs in the short pieces of untranslated sequence amplified by this set of primers.

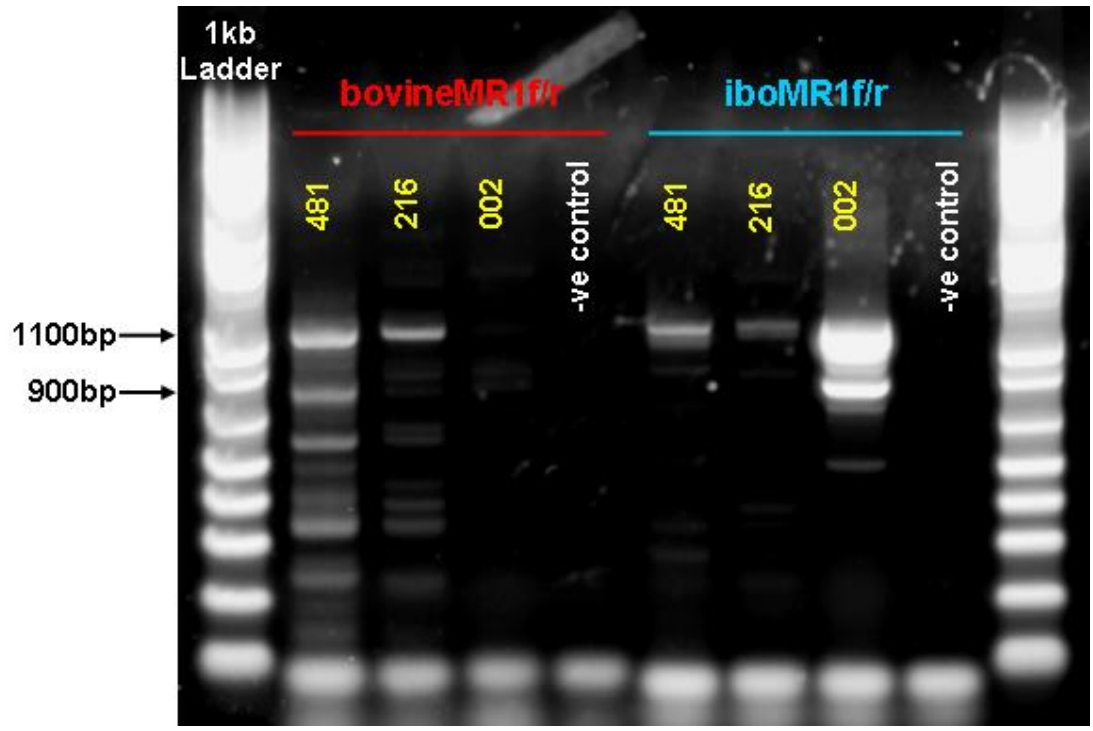


Figure 5-3: Comparison of PCR products obtained from bovine PBMC cDNA using two pairs of primers (bovineMR1f/r and iboMR1f/r) designed to amplify the full-length coding region of bovine MR1. Sterile distilled water was used as a negative control for PCR. A similar result was obtained when the experiment was repeated.

5.3.2 Amplification of multiple ovine MR1 transcripts

PCRs using primer pair bovineMR1f/r failed to amplify full-length products of the predicted size from sheep cDNA. Comparison of human, murine and bovine MR1 sequences demonstrated that the α2 region was most highly-conserved between species. A pair of primers based on conserved segments within the MR1 α2 region was therefore designed and tested on cDNA from ovine PBMC. cDNA from bovine PBMC was used as a positive control. These primers, "boovMR1f2" and "boov(n)MR1r", produced a single product of the anticipated size from both species (see figure 5-4). Sequencing of the products confirmed that the bovine product bore 100% identity to the previously obtained sequence, and the ovine product was 95% identical to bovine, confirming that it represented ovine MR1 α2 region.

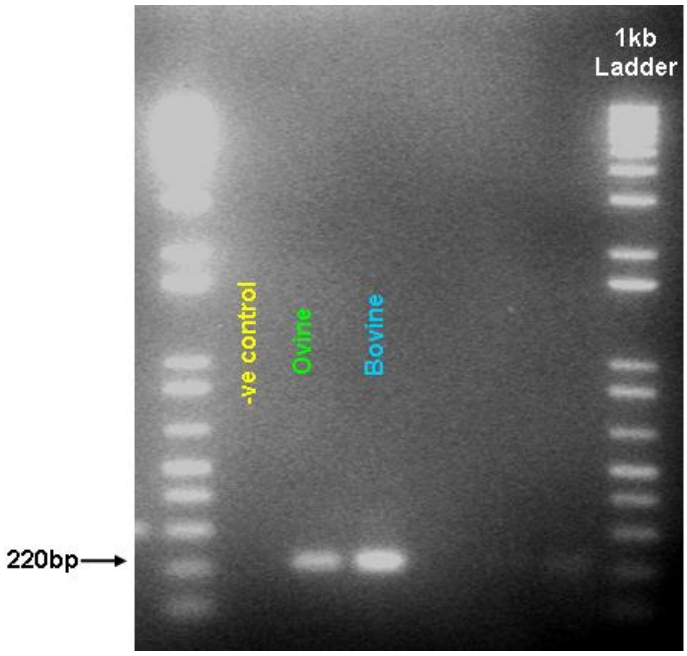


Figure 5-4: Agarose gel showing products amplified from ovine and bovine PBMC cDNA using primers boovMR1f2/boov(n)MR1r located in the α2 region of MR1. These bands were excised and sequenced to confirm that they represented ovine and bovine MR1, respectively. Sterile distilled water was used as a negative control for the PCR.

No *Ovis aries* gene matches to human MR1 nor murine MR1 had been detected in initial searches of gene databases. The partial ovine MR1 α2 sequence, as well as the full-length bovine MR1 sequence, was used to search the ovine gene database using a discontiguous BLAST search on the NCBI website. Three almost identical positive matches were found in *Ovis aries* (EST) (NCBI accession numbers: EE793373, DY502438, EE859755); a high level of similarity to bovine MR1 5' untranslated, leader, α1, α2 and α3 region sequence was demonstrated, although all sequences stopped in the α3 region. To confirm the presence of these sequences in sheep, a forward primer sited in the 5' untranslated region, and reverse primer sited in the α3 exon ("ovMR1f"/"ovMR1r"), were designed from the EST sequences and tested on ovine PBMC cDNA. Following optimisation of the PCR protocol, three bands of 650, 400 and 200 bp were obtained (figure 5-5). Sequencing of the two largest bands

showed that they had >95% similarity to bovine MR1 and were truncated in the α3 region at the site of the reverse primer; the smaller product was missing the α1 region (L, α2, start α3). As with cattle, these data suggested the presence of ovine MR1 splice variants (see section 5.3.5).

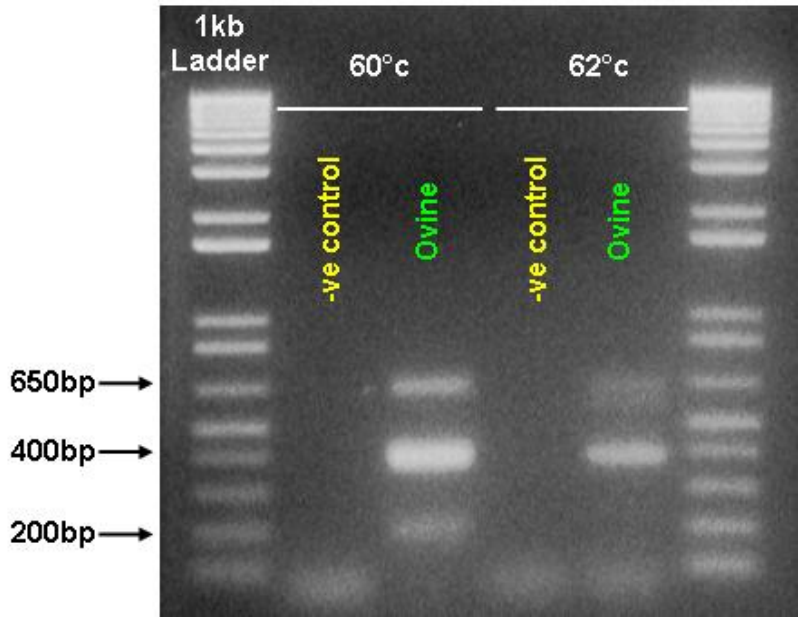


Figure 5-5: Agarose gel of PCR products amplified at two different annealing temperatures from sheep cDNA using primers ovMR1f/ovMR1r located in 5' untranslated and *α*3 regions, respectively, of *Ovis aries* database EST genes identified using BLAST searches of known ovine MR1 *α*2 region sequence. Sterile distilled water was used for PCR negative controls.

At this point in the studies, the bovine primers iboMR1f and iboMR1r, based over the 5' start codon and 3' stop codon of bovine MR1, were made available by Dr Ildiko van Rhijn (see section 5.3.1, above). These primers were tested at a range of annealing temperatures (Tm:60°C-64°C) for their ability to amplify the full-length coding region of ovine MR1 from ovine cDNA. Again, cDNA from bovine PBMC was used as a positive control. Primers iboMR1f and iboMR1r produced two clear bands of 1100bp and 900bp at Tm:60°C, the former corresponding to the expected

full-length product (figure 5-6), similar to those obtained previously from bovine cDNA. Sequencing of cDNA clones derived from DNA extracted from the gels confirmed that the 1100bp product contained the full coding region of ovine MR1, which bore 100% similarity to the previously obtained partial ovine sequences and >94% similarity to bovine MR1. Sequence data from clones containing the 900bp insert were shown to represent an ovine MR1 splice variant missing the α1 exon. The full coding region sequence of ovine MR1 was deposited in the GenBank database of the NCBI (http://www.ncbi.nlm.nih.gov) with the accession number FJ423039. Ovine MR1 sequence data is provided in Appendix D.

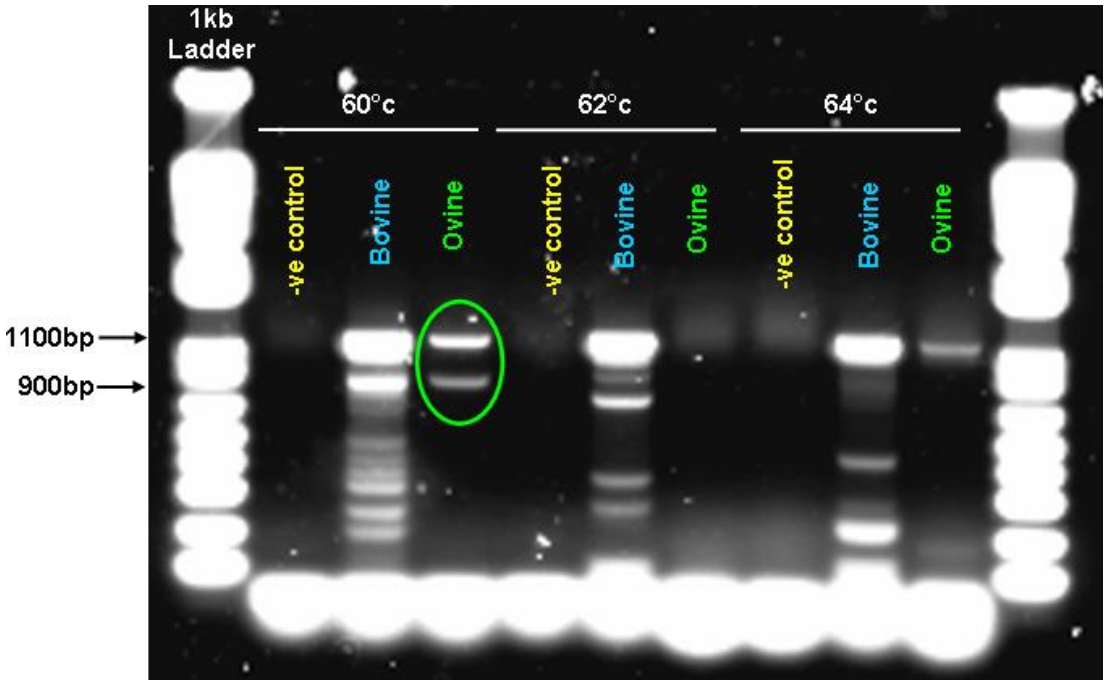


Figure 5-6: Agarose gel showing PCR products amplified at different annealing temperatures from bovine and ovine PBMC cDNA using primers iboMR1f/r. Sterile distilled water was used as a PCR negative control. DNA from the two circled bands of product from ovine PBMC at Tm:60°c were extracted, cloned and sequenced, and confirmed to represent full coding sequence of MR1, and an MR1 splice variant missing the α1 exon.

5.3.3 Identification of genome location and exons of bovine and ovine MR1

In humans and mice, the MR1 gene is located in paralogous regions of the respective genomes in close proximity to other non-classical class I genes. In humans this includes the CD1 family, and the iNKT cell restriction element CD1d. Similarly, while the bovine MHC is located on chromosome 23, searches within the current assembly of the bovine genome (Btau_4.0, available at www.ensembl.org) showed that MR1 sequence maps to chromosome 16 (L: 59794055-109; α1: 59783103-363; α2: 59782169-444; α3: 59780867-1142; TM: 59780158-260). This is sited between the bovine syntaxin 6 gene (Accession number: NP_001075900.1), at 59.71-59.76Mb, which functions in trans-Golgi network vesicle trafficking, and the nuclear protein immediate-early-response 5 gene (IER5) at 59.83Mb. Another identified bovine gene (Acc: IPI00694315), noted as 'similar to xenotropic and polytropic murine leukaemia virus receptor', is sited nearby at 59.56-59.63Mb. Comparison to the corresponding region of both the human genome on chromosome 1q25 and murine genome on chromosome 1G3 shows a similar location of MR1, between the syntaxin 6 and IER5 genes. However, in contrast to humans but similar to mice, the latest assembly of the bovine genome has mapped the bovine CD1 family to a different chromosome, chromosome 3: boCD1a (DQ192541) at 13.13Mb; boCD1b3 (DQ192542) at 12.85Mb; boCD1d psuedogene (DQ192544) at 13.19Mb. Thus, the bovine MR1 gene and the CD1 gene family appear to be located on different chromosomes.

Table 5-2: Genomic locations of the MHC locus and MR1 and CD1 genes in different species.

Class I MHC and MHC-like genes:	Chromosomal Location		
	Human	Murine	Bovine
MHC locus	6	17	23
MR1 gene	1q25	1G3	16
CD1d gene	1q23	3F1	-
CD1 family	1q22-23	3	3

Human and murine MR1 genes demonstrate 'type I splicing' where each domain starts with a composite codon whose first nucleotide is contributed by the previous exon (Riegert et al. 1998). Comparison of bovine MR1 sequence to bovine genomic data enabled identification of distinct MR1 exon sequences, and indicated that these also exhibit type 1 splicing. The different exons of bovine MR1 are shown in Appendix E. Intron/exon boundaries for sheep are likely to be the same as for cattle, because of the high level of similarity between these species' MR1 sequences.

5.3.4 MR1 is highly conserved between species

Full-length sequences of the coding regions of ovine, bovine, human and murine MR1 were compared at the nucleotide and amino acid level; results are shown in table 5-3. The most striking feature of the data is the high level of sequence conservation in the α1, α2 and α3 gene/protein segments (over 75% and up to 97%), which form the extracellular part of the mature protein. In contrast, the leader, transmembrane and cytoplasmic regions show rather low levels of similarity between the species (>55%). Bovine and ovine MR1 sequences exhibited a remarkably high level of nucleotide similarity, which was always >94% in all exons. The α2 region amino acid identity was particularly well conserved at >84% between all species.

Table 5-3: Levels of nucleotide similarity and amino acid identity between species of different segments of the MR1 gene/protein.

	Gene/protein segment					
Sequences Compared	**Leader** 55bp/19aa	**α1** 261bp/87aa	**α2** 276bp/92aa	**α3** 276bp/92aa	**TM** max105bp/35aa	**Cyt** >36bp/12aa
Ov.:Bov.	100%/100%	97% / 97%	95% / 91%	96% / 91%	94% / 88%	95% / 83%
Ov.:Hum.	82% / 74%	84% / 83%	88% / 87%	83% / 73%	79% / 71%	75% / 50%
Ov.:Mur.	76% / 58%	79% / 78%	86% / 87%	75% / 68%	58% / 47%	55% / 25%
Bov.:Hum.	82% / 74%	83% / 82%	88% / 87%	84% / 75%	80% / 71%	73% / 42%
Bov.:Mur.	76% / 58%	78% / 77%	86% / 84%	76% / 72%	58% / 47%	58% / 17%
Hum.:Mur.	73% / 53%	86% / 90%	87% / 89%	82% / 72%	64% / 50%	65% / 25%

Ov: ovine; Bov: bovine; Hum: human; Mur: murine. TM: transmembrane; Cyt: cytoplasmic.

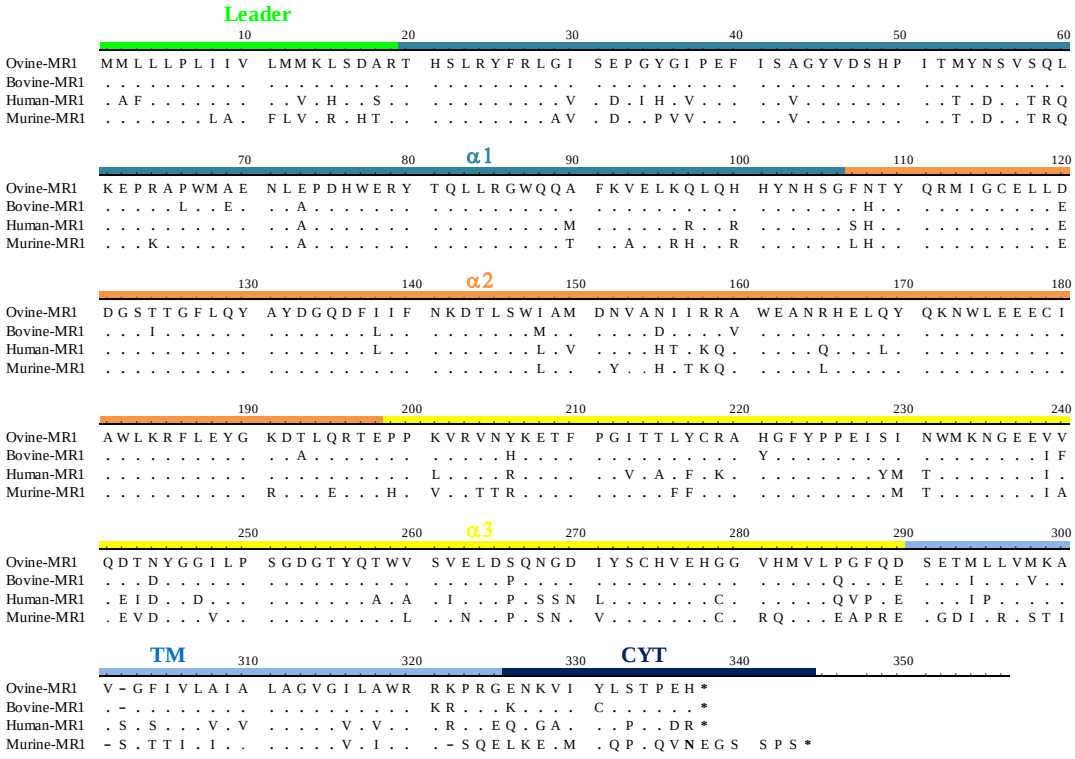


Figure 5-7: Comparison of the amino-acid sequences of MR1 from different species. Different gene/protein segments are indicated; TM: transmembrane; Cyt: cytoplasmic tail. Identical codons between species are identified by a dot, and missing codons by a dash. (Cross-species comparison of full nucleotide sequences are shown in Appendix D).

5.3.5 Analyses of alternatively-spliced variants of MR1

Further analysis was conducted on the splice variants of bovine and sheep MR1 described above, namely the 900bp products from cattle and sheep (named boMR1.900 and ovMR1.900 - NCBI GenBank accession numbers FJ423041 and FJ423042), and the 650bp cattle product (named boMR1.650 - NCBI GenBank accession number FJ423040). These alternatively spliced forms are summarised in figure 5-8 and compared with previously described human and mouse splice variants. Identification of the genomic exon sequences confirmed that the sequences missing from boMR1.900, ovMR1.900 and boMR1.650 corresponded exactly to the complete

α1 exon for the former and the α1 and α3 exons for the smaller cDNA. The corresponding ends of bordering exons were joined by type 1 splicing, as for the full length gene. These sequences did not contain internal stop codons and retained the original reading frame. Hence, they potentially can generate truncated forms of the MR1 protein. Full nucleotide sequences for each bovine and ovine splice variant are provided in Appendix F.

Following the initial discovery of multiple MR1 transcripts in bovine PBMC, the existence of MR1 splice variants was confirmed by testing reverse and forward primers based in the α2 exon (boMR1a2-r1/r2 and boMR1a2-f1/f2), in combination with a forward primer in the 5' untranslated region (bovineMR1f) or a reverse primer in the 3' untranslated region (bovineMR1r), on bovine PBMC cDNA. Each PCR produced two products of anticipated size (600 and 300bp, and 700 and 400bp, respectively), representing full MR1 and a form missing either the α1 or α3 exon (data not shown).

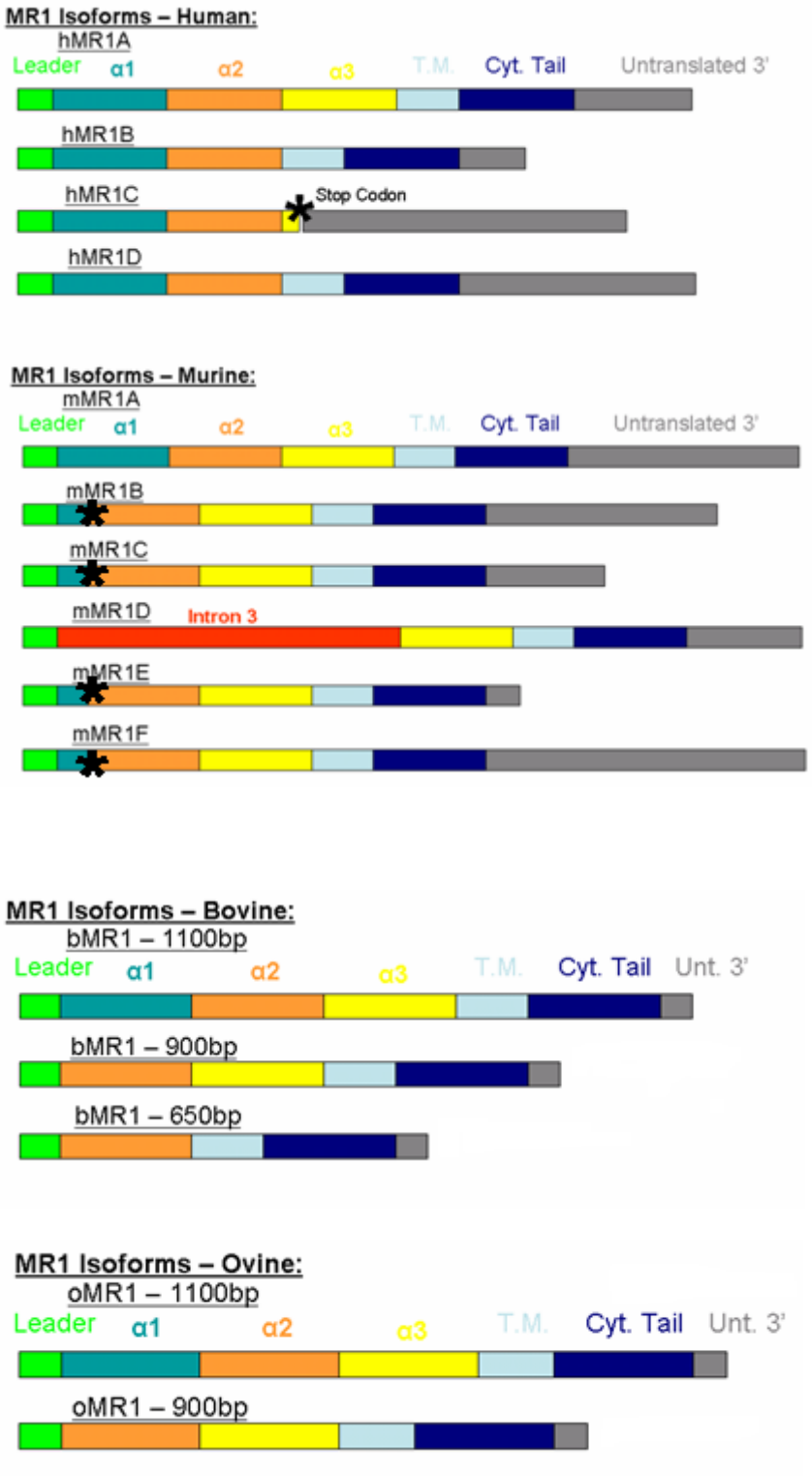


Figure 5-8: Diagrammatic representation of alternatively spliced transcripts of bovine and ovine MR1 compared with splice variants previously described in humans and mice (Riegert et al. 1998). The presence of a stop codon is indicated by an asterisk.

Additional results have demonstrated the potential for other splice variants of MR1. In experiments to study MR1 expression (see chapter 6), cloning of full-length bovine MR1 into vector, and sequencing of multiple inserts, revealed one clone (from four) with missing sequence in the α3 exon. A discrete 87bp of sequence extending from the 5' intron/exon junction of α3 to the middle of the α3 exon (bp593 to bp680) was excluded, although sequencing data on either side was 100% identical to the anticipated full MR1 sequence (see Appendix G for nucleotide sequence). Sequencing of the plasmid DNA was repeated on two occasions with matching results. This deletion also did not alter the open reading frame of the sequence, or result in any internal stop codons, suggesting the ability to produce potentially functional protein.

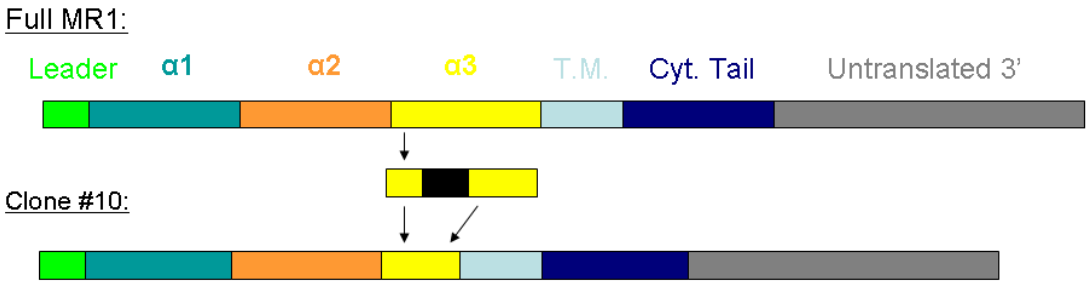


Figure 5-9: A further splice variant of bovine MR1 identified by subcloning of cDNA thought to represent full MR1 transcripts. Clone (#10) contained an MR1 transcript with a shortened α3 region, missing a 90bp sequence, but retaining an open reading frame.

5.4 Discussion

Orthologous MR1 genes have been detected in a number of species, including humans, primates, mice and rats. The work described in this chapter identified the full sequences for the coding regions of bovine and ovine MR1, and demonstrated a high level of sequence conservation between species. A number of abundant alternatively-spliced variants of MR1 were also detected in both cattle and sheep. Additionally, the genomic location of MR1 was identified as being within a paralogous region of the bovine genome but on a different chromosome than CD1.

Prior to this study, MR1 had not been identified in ruminants. Human and murine MR1 sequences were used to search online bovine and ovine sequence databases and, using a combination of EST and genomic partial sequences, PCR primers were designed to amplify fragments and eventually the full coding regions of bovine and ovine MR1. These data confirmed that MR1 is highly conserved between mammalian species.

Bovine MR1 was mapped to *Bos taurus* chromosome 16 (BTA16). Although the MR1 sequence in the genome is incomplete, the available data indicate that it has an intron-exon structure similar to that seen in other species, including a large intron between the first two exons (Yamaguchi et al. 1998). The genomic location of bovine MR1 and the identity of the adjacent genes indicate that it occupies a region paralogous to the genomic regions of mice and humans containing MR1. Additionally, the upper third of BTA16 has been shown to correspond to murine chromosome 1, where murine MR1 is located (Schmutz et al. 2001). Although a majority of the gene sequences located near the site of MR1 on BTA16 have yet to be fully identified, the adjacent genes to MR1 were identified as syntaxin 6 and IER5, and these are also the adjacent genes to MR1 on human chromosome 1q25 and murine chromosome 1G3.

There is evidence that class I MHC-like genes in humans are clustered in at least two of four MHC paralogous regions of the genome and it has been proposed that these

represent ancient duplication events (Shiina et al. 2001; Abi-Rached et al. 2002). The region (1q21-25) of human chromosome 1 which contains MR1 and the CD1 family of genes has been shown to contain a significant number of other immunological genes, including SPTA1, MNDA, IFI-16, AIM2, BL1A, FY and FCERIA (Shiina et al. 2001). It is interesting to note that the majority of these react with lipids/glycolipids, as has been hypothesised for MR1 (Shimamura et al. 2007), prompting the suggestion that the genes within this region represent a gene cluster dedicated to processing and presentation of lipid/glycolipid antigens.

BTA16 has been identified to contain four evolutionary breakpoint regions (EBR) which are specific to the ferungulata cohort, suggesting that the chromosomal regions were rearranged before divergence of the Artiodactyla and Carnivora, and thus have probably been retained in the ancestral form within the human genome (The Bovine Genome Sequencing and Analysis Consortium et al. 2009). However, although both MR1 and the CD1 family of genes both map to the paralogous region of human chromosome 1, the identified bovine CD1 genes map to a different area of the *Bos taurus* genome from MR1, at chromosome 3. This is similar to the situation in mice where MR1 is located on chromosome 1 but the CD1 family of genes is located on chromosome 3 (Moseley et al. 1989). This contradicts the proposal that these genes remain closely linked because of their functional relatedness. The separate chromosomal location of the bovine CD1 genes confirms that further breakpoints have occurred in bovine chromosome 16 subsequent to divergence of Artiodactyla and Carnivora (as suggested by Womack and Moll 1986; Womack and Kata 1995).

Comparison of MR1 coding sequences identified a high level of sequence conservation between species. The combined α1 and α2 domains of bovine MR1 showed 81% and 85% amino acid identity to murine and human MR1 respectively; the murine and human sequences were 90% identical. This compares with 73% identity for the corresponding domains of bovine and human CD1b, and 61% for murine and human CD1d. Overall, the most striking aspect of these comparisons was the high level of conservation of the sequences encoding the extracellular

domains of the protein (α1, α2 and α3), in comparison with the leader, transmembrane and cytoplasmic regions, which in some cases show very low levels of conservation. This suggests evolutionary pressure to conserve the extracellular structure of the protein.

Although the way in which MAIT cells interact with MR1 has not been defined, the likely interaction can be inferred from knowledge of how conventional T-cells recognise classical class I MHC, where the CDR loops of the TCR chains interact with the α-helices of the peptide binding groove and with residues in the bound peptide protruding from the groove (Garboczi et al. 1996; Garcia et al. 1998). The α1 and α2 domains make up the binding groove, and these are highly polymorphic in classical class I proteins to provide the capacity to bind a large variety of TCRs. The high degree of conservation of MR1 between species suggests that it may bind a conserved ligand, or set of ligands. Currently, the chemical nature of the MR1 ligand is unknown. It is likely that, like other ligand binding class Ib molecules, MR1 binds both exogenous and endogenous ligands; however, although MAIT cells have been shown to be activated *in vitro* by a synthetic glycolipid (Okamoto et al. 2005), molecular modelling of the α1/α2 regions suggests that they are hydrophilic and more similar to classical class I molecules than the lipid-binding CD1 molecules (Riegert et al. 1998).

Cross-species conservation of MR1 is consistent with the highly-conserved nature of the CDR sequences of the MAIT TCRα chain between species. During antigen-presentation to conventional αβ T cells, the α2 domain of classical class I MHC molecules interacts with the α chain of the cognate TCR, in a process which mediates the interaction and orientation of TCR-MHC binding (Garboczi and Biddison 1999; Garcia et al. 1999). If the MAIT TCR interaction with MR1 adopts a similar orientation, the MR1 α1 domain, which although less so is also highly conserved between species, would be expected to interact with the CDR1and CDR2 of the MAIT TCRβ. MAIT TCRs have an oligoclonal Vβ repertoire and utilise predominantly Vβ2 or Vβ13 in humans, and Vβ6 or Vβ8 in mice; however, MAIT TCRVβ chains show no restriction in Jβ usage or CDR3 length (Tilloy et al. 1999).

The restricted nature of the MAIT Vβ repertoire may be important with regard to the ability of MAIT cells to interact with the MR1 α1 domain. Also, by inference from TCR-classical class I structure, diversity in the CDR3 of the β chain would allow some flexibility for recognition of different MR1-bound ligands. The MR1 α3 domain is not involved in ligand binding but interacts with β2-microglobulin. Hence, there may be less selective pressure for conservation of this region, and indeed, since β2M shows some variation in sequence between species, some of the variation in α3 sequence may be related to differences in the sequence of β2M.

A striking feature of the analysis of MR1 transcripts in both cattle and sheep was the detection of an alternatively-spliced variant lacking the α1 exon; this was highly abundant both in PBMC and in several bovine cell lines (see section 6.3.1). A second less abundant transcript, which lacked both the α1 and α3 domains, was also identified in cattle. Confirmation of MR1 splice variant structure was made through use of additional PCR primers based within the α2 exon. A number of alternatively spliced variants of MR1 have also been described in humans and mice (Riegert et al. 1998). Although many of the splice variants in the mouse also had disrupted α1 domains, these mouse variants retained a small part of the α1 exon; however, these all contained aborted open-reading frames. By contrast, the two human MR1 variants which maintained an open-reading frame exhibited omission of the α3 exon. This current study also revealed an additional cattle transcript with in-frame disruption of α3 domain sequence, which may represent a further potential splice variant of bovine MR1. A number of the human and murine splice variants display differences in the length of the 3' untranslated region sequence, and preliminary data in cattle suggests that further splice variants may also exist in this species dependent on 5' untranslated sequence length.

A consistent feature of the abundant MR1 splice variants is that they retain an intact α2 exon. Both of the ruminant splice variants retained the original reading frame and would be predicted to encode polypeptides with intact transmembrane and cytoplasmic domains, although they are unlikely to retain any function in relation to presentation to MAIT cells. However, given the abundance of the α1-deleted

transcript in ruminants and the evidence that surface expression of MR1 may be tightly regulated, it is possible that the product of this splice variant is somehow involved in regulating surface expression of the intact protein. Current data indicate that MR1 is transported through the endoplasmic reticulum and Golgi apparatus to endosomal compartments, where it has been suggested that it acquires a ligand (Huang et al. 2008). If indeed the splice variants have some functional role, the consistent retention of α2 may suggest that this domain is required for endosomal targeting.

6 Analysis of Bovine MR1 Expression

6.1 Introduction

MR1 shows similarity in sequence and protein structure to classical class I MHC proteins. The extra-cellular portion of the MR1 heavy chain consists of three domains, α1, α2 and α3, two of which (α1, α2) form a groove bounded by α-helices into which a ligand is believed to bind. The α3 domain, which adopts an immunoglobulin fold structure, associates with the light chain, beta-2-microglobulin (β2M). Analysis of MR1 RNA expression by Northern blotting and rtPCR has demonstrated ubiquitous expression in a variety of human and murine tissues, as well as in human cell-lines of different lineages. Full-length as well as truncated alternatively spliced transcripts of MR1 were observed in all cell-lines and tissues tested (Hashimoto et al. 1995; Yamaguchi et al. 1997; Riegert et al. 1998).

Despite the ubiquitous expression of MR1 transcript, cell surface expression of the protein has not been demonstrated *in vivo*, and has been difficult to achieve in cells transfected with the MR1 heavy chain *in vitro*. Cell surface expression has been shown to be dependent on β2M (Yamaguchi and Hashimoto 2002; Huang et al. 2009). However, sequence analysis has suggested that MR1 lacks the phylogenetically-conserved motifs implicated in β2M association with classical class I molecules, and therefore is predicted to have a weaker association with β2M (Miley et al. 2003). Most cell-lines transfected with MR1 cDNA express the protein intracellularly, but low levels or in some cases no MR1 is expressed on the cell surface. Recent evidence shows that, like CD1d, MR1 is transported from the Golgi to the endosomal compartment, and only limited amounts of the protein reach the cell surface from there (Huang et al. 2008). Hence, these observations suggest that MR1 is dependent on endogenous ligand for surface expression and that this may be acquired within the endosomal compartment. Since peripheral expansion of MAIT cells is dependent on the presence of a commensal gut flora (Treiner et al. 2003), it has been hypothesised that gut micro-organisms may either provide exogenous

ligands or induce endogenous ligands for MR1, or induce translocation of MR1 to the cell surface via a cellular 'stress' signal (Hansen et al. 2007).

Previous studies of cells transfected with the murine MR1 heavy chain showed that replacement of the α3, transmembrane (TM) and cytoplasmic (Cyt) domains of murine MR1 with the corresponding domains of a classical class I molecule (murine L^d) increased levels of cell-surface expression (positive cells increasing from 10% to 26% of the transfected mouse L cells) (Miley et al. 2003).

The aims of the studies described in this chapter were, firstly to determine whether the levels and type of MR1 transcript expressed varies between bovine cell lines belonging to different leucocyte lineages, and secondly to prepare constructs to study in vitro expression of MR1 in transfected cells. The latter was undertaken to provide eukaryotically expressed protein to examine potential cross-reactivity of antibodies raised against MR1 in other species and with the eventual aim of investigating the factors that influence cell surface expression of the protein.

6.2 Materials and Methods

6.2.1 Bovine cell lines and PCR for MR1 transcripts

Generation and maintenance of cell-lines are described in section 2.3.1.4. Cell lines used were: BL20, a lymphoma cell line believed to be of B-cell origin; 592TPM, a B-cell line infected with *T. parva*; 663TA/B, a B-cell line infected with *T. annulata*; 592TAA and 663TA, monocyte cell lines infected with *T. annulata*; 109TPM, a CD8 T-cell line, 158TPM, a T-cell line of mixed CD8/γδ phenotype, and 951TPM, a γδ T-cell line, were all infected with *T. parva*. A further two cell lines of undefined phenotype infected with *T. annulata* and *T. parva* respectively, were also used. RNA was extracted from cell-lines using Tri-Reagent (see section 2.4.1.1) and reverse-transcribed to cDNA (section 2.4.2.1). The PCR protocol for detection of MR1 transcripts in cell-lines was as described in section 2.4.2.2: this utilised primer pairs "bovineMR1f" / "bovineMR1r" and "iboMR1f" / "iboMR1r", using melting temperatures of 58°C and 60°C, respectively. Products were analysed by 1.5% agarose gel electrophoresis (section 2.4.3.3).

6.2.2 Classical MHC class I in pcDNA3.1 / V5-His TOPO TA expression vector

Plasmid DNA from a bovine classical MHC class I (N*02401) transformed into the TOPO TA expression vector was provided by Dr. Jane Hart, University of Edinburgh. This had been previously demonstrated to be transfectable and recognised by a number of class I-specific monoclonal antibodies (Ellis et al. 1996), and was used as a positive control for the transfection reaction and analysis of MR1 protein expression.

6.2.3 Preparation of cDNA constructs for incorporation into the pFlag-CMV-3 vector

Full MR1 coding sequence, as well as a chimaera of MR1 α1 and α2 regions in combination with the α3, TM and Cyt regions of a classical MHC class I, were ligated into a pFlag-CMV-3 vector. This vector was chosen for transfection studies as it contains a sequence encoding a peptide tag at the N–terminus, allowing detection of cell surface-expressed protein with an anti-tag antibody (M2). The vector also contains a methionine start codon and 'leader' sequence, and therefore primers were designed to omit this region from each insert gene during PCR amplification.

6.2.3.1 Preparation of the construct containing the full coding region of MR1

The full MR1 coding sequence (without the signal sequence) was amplified from cDNA of bovine PBMC from animal #332, with primers "fMR1f" and "fMR1r" (see table 6-1), using standard PCR reaction mixture (see section 2.4.2.2) but cycling conditions of 95°C for 10 minutes, 40 cycles of (1 minute at 95°C, 1 minute at 60°C, 1 minute at 72°C), 10 minutes at 72°C. These primers were designed to add HindIII and XbaI restriction sites to the beginning and end, respectively, of the amplified MR1 sequence.

6.2.3.2 Preparation of the class I/MR1 chimaeric construct

Bovine PBMC from an animal (cow 641) known to be homozygous for the haplotype carrying the N*01301 class I gene, which encodes the A18 serological specificity, were used as a source of cDNA to isolate the relevant segment of the N*01301 gene. The full coding sequence was initially isolated by PCR using primers "Bo7" and "Bo3-1_rev" specific for classical MHC class I sequence (see table 6-1; designed by Dr. Niall MacHugh, University of Edinburgh), employing standard reaction mixtures and cycling protocol of 95°C for 3 minutes, 5 cycles of (95°C for 1 minute, 65°C for 1 minute, 72°C for 2 minutes), 25 cycles of (95°C for 1 minute, 55°C for 1 minute,

72°C for 2 minutes), 72°C for 10 minutes. The PCR product was purified and subcloned into pGEM-T Easy vector (see section 2.4.4), and the N*01301 sequence was confirmed by sequencing at DBS genomics (section 2.4.3.6).

The original bovine N*01301 and MR1 sequences were checked to confirm that they did not contain the BsiWI restriction site. A partial MHC sequence containing the α3, TM and Cyt regions was PCR amplified from N*01301 DNA, with primers "cMHCf" and "cMHCr" (see table 6-1), using a standard PCR protocol (see section 2.4.2.2) and Tm:70°C. These primers were designed to add restriction enzyme sites BsiWI immediately before the start of the α3 region, and XBaI after the stop codon in the cytoplasmic tail. A partial MR1 sequence containing the α1 and α2 regions was PCR amplified from bovine DNA using primers "cMR1f" and "cMR1r" (see table 6-1) using cycling conditions of 95°C for 10 minutes, 40 cycles of (1 minute at 95°C, 1 minute at 62°C, 1 minute at 72°C), 10 minutes at 72°C, and standard reaction mixtures. These primers added a HindIII restriction site before the α1 region and a BsiWI site immediately after the end of the α2 region.

Table 6-1: Primers used in PCR and optimal annealing temperature (Tm) for each reaction.

Target Sequence	Forward 5' sequence 3'	Reverse 5' sequence 3'	Optimal Tm
Full MR1 (pFlag vector)	fMR1f AGTCTCAAGCTTCGGACTCACTCTC TGAGATATTTTCG	fMR1r GGAGATGGCTGCTTCCACCTTTCTAGAC TATGA	60°C
Chimaera MR1 (pFlag vector)	cMR1f TCAGCTAAGCTTCGGACTCACTCTC TGAGATATTTTCG	cMR1r GGAAAGATGCCTTACAAAGAACACGTA CGACGTGA	62°C
N*01301 (MHC A18)	Bo7 GGCTACGTGGACGACACG	Bo3-1_rev CTGAGTGATGCTTCATCC	65/55°C
Chimaera MHC (pFlag vector)	cMHCf ATGACTCGTACGGACCCTCCAAAGG CACATGTG	cMHCr GACAACTGCCTTGTGGGGACTGTCTAG ACTGCAA	70°C

Target sequences were – Full MR1: MR1 coding sequence without the leader sequence; Chimaera MR1: α1 and α2 regions of MR1; N*01301: Full coding sequence of the N*01301 gene; Chimaera MHC: α3, TM and Cyt regions of the N*01301 gene.

6.2.4 Transformation of the MR1 and the MR1/MHC chimaera constructs into the pFlag-CMV-3 expression vector

Products of predicted size obtained from PCR reactions were excised from agarose gels and further amplified by PCR, using the same primers and PCR protocols, and the DNA was purified (see section 2.4.3.5). This additional step was to increase concentration and purity of product without the use of subcloning, since additional vector sequence and nucleotide fragments may have reduced the success of subsequent reactions. Consensus DNA sequences were confirmed by direct sequencing of purified PCR product. Vector and the full length coding sequence of MR1 cDNA were digested using HindIII and XbaI; the Chimaera MR1 DNA was digested using HindIII and BsiWI; Chimaera MHC DNA was digested using BsiWI and XbaI (see section 2.4.1.5).

Ligation reactions, transformation into JM109 competent cells, and PCR-screening of colonies for inserts are covered in sections 2.4.4 and 2.4.5. Positive colonies were expanded overnight in LB medium and plasmid DNA extracted by either MiniPrep or MaxiPrep reactions (section 2.4.1.3). Insert DNA was sequenced by DBS Genomics, Durham University, UK (section 2.4.3.6) and assessed for correct sequence and an intact open reading frame using DNASIS software (Hitachi Software Engineering Company, Ltd.) (see section 2.5).

6.2.5 Transfection into Cos-7 cells

Cos-7 cells were maintained and passaged as described in section 2.3.2.1. Transfection of plasmid DNA was performed using FuGENE HD Transfection Reagent (Roche Applied Science, Burgess Hill, UK) (see section 2.3.2.2).

6.2.6 Analysis of protein expression

Monoclonal antibodies used for immunofluorescence staining of transfected cells, and their specificities, are shown in table 6-2 (see also section 2.3.3.1, table 2-1). Transfected Cos-7 cells were assessed for cell surface expression using flow cytometry analysis (as described in section 2.3.3.2), and intracellular expression was examined by fluorescent microscopy of cytospin preparations and in-well adherent cells, which were fixed prior to staining with acetone or acetone/ethanol (see section 2.3.3.4).

Table 6-2: Specificities of monoclonal antibodies used for primary antibody staining for protein expression analyses.

Antibody	Specificity
IL-A88	Bovine MHC Ia heavy chain
IL-A19	Bovine MHC Ia in association with β2M
MR1 26.5	Human MR1
W6-32	Human MHC Ia and some bovine MHC Ia alleles
JG-9	Bovine β2M
BG10	Bovine β2M
M2	pFlag-CMV-3 Vector tag

Aliquots of monoclonal antibody 26.5 in the form of hybridoma supernatant were provided by Ted Hansen, Washington University School of Medicine, Missouri, USA. Supernatant was used undiluted in all analyses. Antibody 26.5 had been produced in mice against soluble human MR1/β2M molecules secreted from transfected insect cells (Miley et al. 2003; Huang et al. 2005). This antibody had previously been reported to detect, with varying affinity, surface expression of transduced mouse, rat, human and bovine MR1 on human and murine cell lines, following their transduction into these cell lines using a recombinant retrovirus system (Huang et al. 2009).

6.3 Results

6.3.1 Bovine MR1 is abundantly transcribed in different bovine cell lineages

The protozoan parasites *T. parva* and *T. annulata* infect and transform different lineages of leucocytes, the former T and B lymphocytes and the latter B lymphocytes and monocytes. The availability of a series of Theileria-transformed cell lines representing these different lineages, including different subsets of T cells, provided an opportunity to investigate whether expression of MR1 differed according to cell lineage. Transcription of MR1 was assessed by PCR. Initial experiments utilised single *T. parva*-transformed and *T. annulata*-transformed lines of unknown phenotype and a bovine lymphoma cell line, BL-20. Initial testing of two primer pairs "bovineMR1f/r" and "iboMR1f/r", which had had previously been shown to give different results with PBMC from different cattle, showed that the primers "bovineMR1f/r" produced only weak product and demonstrated poor amplification in all reactions. However, primers "iboMR1f/r" produced abundant product of the expected size for the full MR1 coding region from all 3 cell lines and PBMC cDNA (Figure 6-1); a weaker band corresponding in size to the previously described MR1 splice variant missing the α1 exon was observed in the 3 cell-lines, but not in the PBMC cDNA.

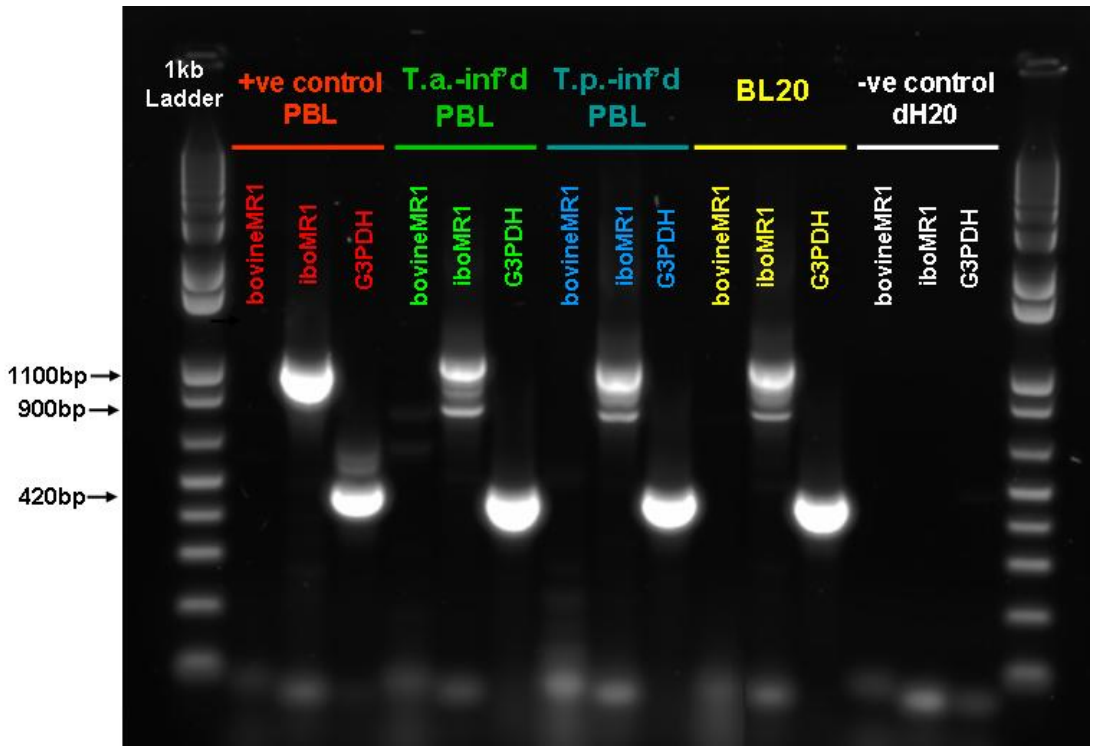


Figure 6-1: Expression of multiple transcripts of MR1 (1100bp, 900bp) in two cell-lines of unknown phenotype derived from bovine PBL and infected with either *Theileria parva* or *T. annulata* (T.p.-inf'd and T.a.-inf'd, respectively), and a bovine lymphoma cell-line (BL20), in comparison to PBL, and a PCR negative control of sterile distilled water. Two different primer pairs ("bovineMR1f/r" and "iboMR1f/r") were used to amplify MR1, and assessed in comparison to the amplification of the housekeeping gene, G3PDH (420bp).

To investigate the effect of cell lineage on MR1 expression, seven cell lines transformed with *T. parva* or *T. annulata*, representing T cell, B cell and monocyte lineages, were examined by PCR using primers "iboMR1f/r". The results shown in figure 6-2 demonstrated abundant amplification of the full coding region of MR1 product from all seven cell-lines, and less abundant amplification of smaller product corresponding to the MR1 splice variant lacking α1 (see figure 6-2). Amplification was performed using equal amounts of cell-line cDNA, and was compared to the housekeeping gene, G3PDH. MR1 product amplified from B-cell lines was more abundant than that from monocyte or T cell lines. There was no apparent association with the animal from which the cell line was derived, since the same two animals

were each the source of a B cell line and a monocyte cell line. Moreover, the three T cell lines, which were derived from different animals (and represented different T cell lineages), gave similar PCR products.

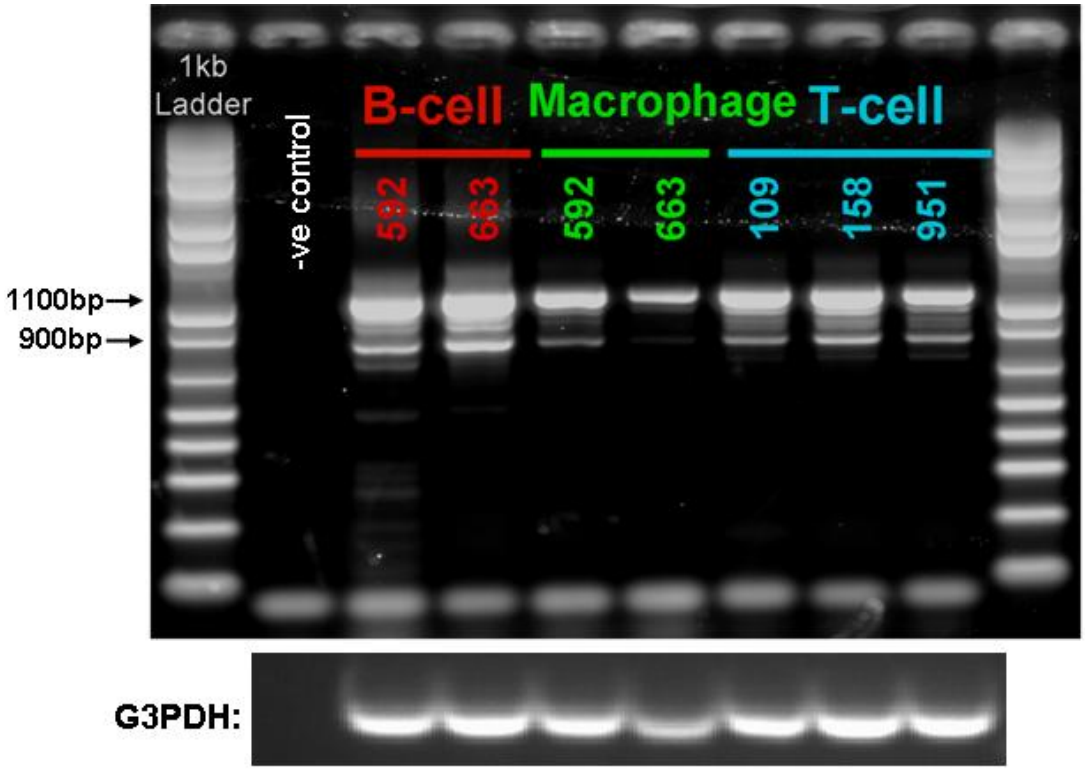


Figure 6-2: Ubiquitous expression of full MR1, and an MR1 splice variant, in specific bovine leucocyte cell-lines infected with *Theileria spp.*. B-cell and monocyte cell-lines were isolated from the same two animals, 592 and 663; T-cell lines were phenotyped by flow cytometry as pure $CD8^+$ (109), pure $\gamma\delta^+$ (951), or mixed $CD8^+/\gamma\delta^+$ (158) T-cell subpopulations. MR1 was amplified using primers "iboMR1f/r", and assessed in comparison to amplification of the housekeeping gene, G3PDH. A PCR negative control of sterile distilled water was used.

6.3.2 Generation of MR1 and MR1/MHC chimaeric constructs for incorporation into a 5'-tagged expression vector

The pFlag-CMV-3 expression vector was used to study expression of MR1 in transfected cells because it encodes a tag on the 5' end of the protein that allows both intracellular and cell surface detection of expressed protein using a tag-specific antibody (M2). Since previous studies in mice showed that grafting of the α3, transmembrane and cytoplasmic domains of a classical class I gene onto the α1 and α2 domains of MR1 enhanced cell surface expression, a similar chimaeric construct was produced for the bovine gene. The monoclonal antibody IL-A88 is specific for bovine classical class I heavy chain; hence, it was anticipated that this antibody might also be used to detect expression of the chimaeric protein. An expression construct containing a classical bovine class I gene within a TOPO vector was also used as a positive control.

The full length coding sequence for MR1, missing the signal sequence, was cloned, sequenced and inserted into the pFlag-CMV-3 vector as described in sections 6.2.3 and 6.2.4. The previously described protocol for preparing a murine MR1 chimaeric gene involved creating a BsiWI restriction enzyme site for ligating the MR1 and MHC segments. Following this protocol, PCR primers for amplifying the bovine MR1 and MHC segments were designed to generate enzyme restriction sites and to retain the original amino acid sequences, but with the addition of two novel amino acids at the junction of the two segments (representing the BsiWI site). Since both MR1 and classical class I molecules undergo type-I intron/exon splicing, where each domain starts with a composite codon whose first nucleotide is contributed by the previous exon, to avoid altering the original amino acid sequence around the BsiWI site it was necessary to remove a guanine nucleotide from the 3' end of the MR1 α2 region, and include this at the start of the N*01301 α3 region.

Products of the expected sizes were obtained using these primers, and consensus DNA sequences were confirmed by direct sequencing of purified PCR product, before the inserts were excised by digestion with HindIII, XbaI and BsiWI enzymes.

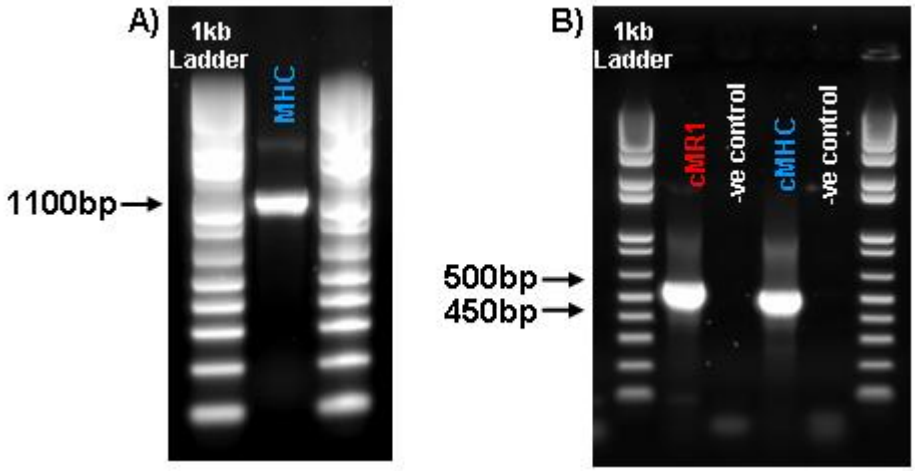


Figure 6-3: Production of DNA segments of MR1 α1 and α2 regions (cMR1), and the α3, trans-membrane and cytoplasmic regions of a classical class I MHC, N*01301 (cMHC). A) Full-length N*01301 cDNA (MHC) amplified from bovine PBMC cDNA. cDNA used was from an animal known to be homozygous for the MHC haplotype carrying this gene. B) Amplification of cMR1 regions from bovine PBMC cDNA, and of cMHC regions from a subclone of the N*01301 gene obtained in gel A. DNA was extracted from the bands of the correct size and re-amplified by the same primers, and PCR product consensus DNA was confirmed by sequencing.

The MR1 and N*01301 segments were added together and ligated into the pFlag-CMV-3 expression vector in the same reaction, and the ligation product was transformed into JM109 competent cells. PCR analysis of 39 colonies with vector-specific primers showed inserts of the size expected of a chimaeric-insert in 31 colonies. Sequencing of the inserts in 12 of these colonies confirmed that in 5 clones the two components of the insert were in the correct order, that they had an intact open reading frame and had correct nucleotide sequences (see figure 6-4). The remaining 7 clones showed one or two single nucleotide substitutions, possibly as a result of PCR errors.

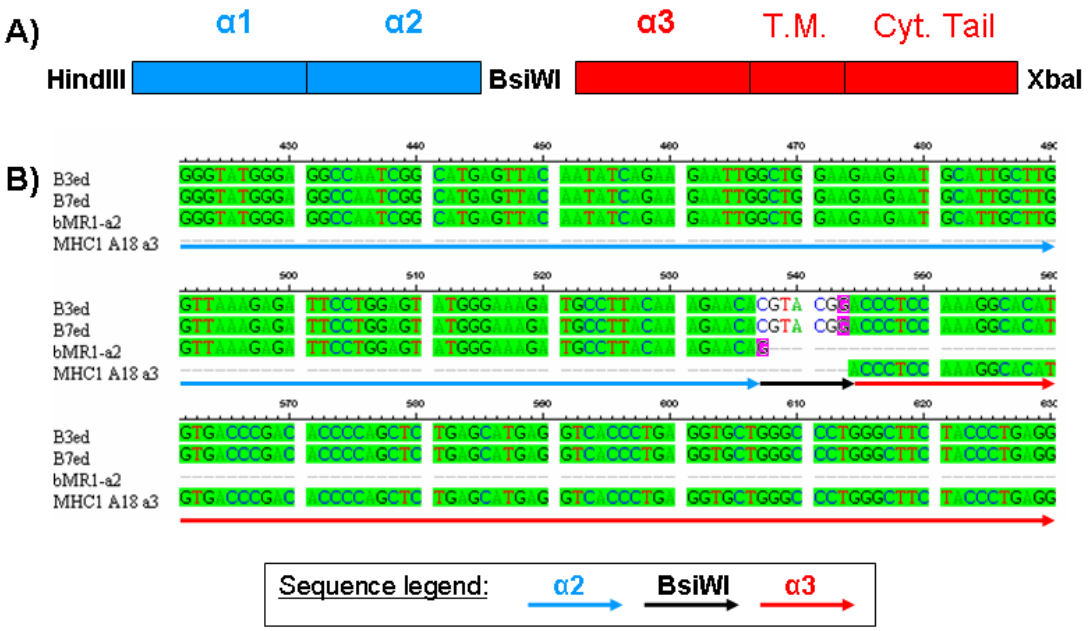


Figure 6-4: A) Schematic representation of MR1/MHC chimaeric sequence (blue:MR1; red:MHC N*01301) and endonuclease restriction sites employed for annealing the fragments (BsiWI) and ligation into the vector (HindIII, XbaI). B). Confirmation of the sequence of two MR1/MHC chimaera clones, B3ed and B7ed, and comparison with sequences of the fragments used to create the chimera, bMR1-a2 and MHC1 A18 a3. Green highlighting indicates an exact match to the MR1α2 and MHCα3 sequences. The two codons added by creating the BsiWI restriction site sequence CGTACG are shown (unhighlighted) at the junction of α2 and α3. To ensure that this sequence did not affect the original reading frame, and since the original molecules both undergo type-I splicing where each domain starts with a composite codon whose first nucleotide is contributed by the previous exon, the guanine nucleotide at the 3' end of bMR1-a2 sequence (highlighted pink) has been moved to the 5' start of the chimaera α3 exon after the site of BsiWI.

6.3.3 Transfected MR1 is expressed intracellularly but not on the cell surface

Cos-7 cells transfected with a classical class I gene or MR1 or the chimaeric construct, were stained by immunofluorescence, either as live cell suspensions or as fixed cells on slides or in culture plates. Both types of preparation were analysed with four monoclonal antibodies: M2, specific for the vector 5' tag, 26.5 raised against human MR1, IL-A88 specific for bovine classical MHC class I heavy chain, and IL-A19 specific for bovine MHC class I in combination with β2M. Transfected cells analysed by flow cytometry were also assessed with a further 3 antibodies: W6-32, specific for human MHC class I heavy chain and cross reactive with some alleles of bovine MHC class I, and JG9 and BG10 both raised against bovine β2M (see section 2.3.3.1, table 2-1).

Since Theileria-infected cell lines were found to contain abundant MR1 transcripts, one of these lines (592TPM) was initially tested by flow cytometry for surface staining with antibody 26.5 raised against human MR1. Antibodies IL-A88 and IL-A19 were also tested and both were found to give strong staining of the majority of cells. By contrast no surface staining was detected with antibody 26.5 (see figure 6-5).

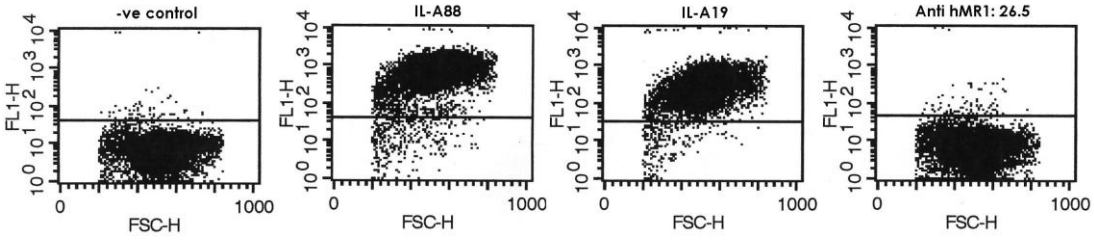


Figure 6-5: Flow cytometry analyses of a Theileria-transformed cell-line stained with primary antibodies specific for bovine class I MHC heavy chain (IL-A88), bovine class I MHC heavy chain in association with β2M (IL-A19), human MR1 (26.5), and an immunofluorescent secondary antibody. A negative control was incubated with FACS medium and secondary antibody.

Analysis by flow cytometry of untransfected Cos-7 cells stained with the different antibodies (as a negative control) confirmed negative results for M2, 26.5 and IL-A88. However, the cells were weakly positive for IL-A19 (16%), and strongly positive for W6-32 (99%), JG9 (99%) and BG10 (95%) (see figure 6-6). The intensity of staining with these antibodies on cells transfected with MR1 and the chimaeric MR1 was similar to that obtained with the untransfected cells. These results suggest that bovine β2M within the bovine calf serum used in the cultures was interacting with simian MHC class I molecules on the Cos-7 cells, and that W6-32 was able to detect simian class I MHC.

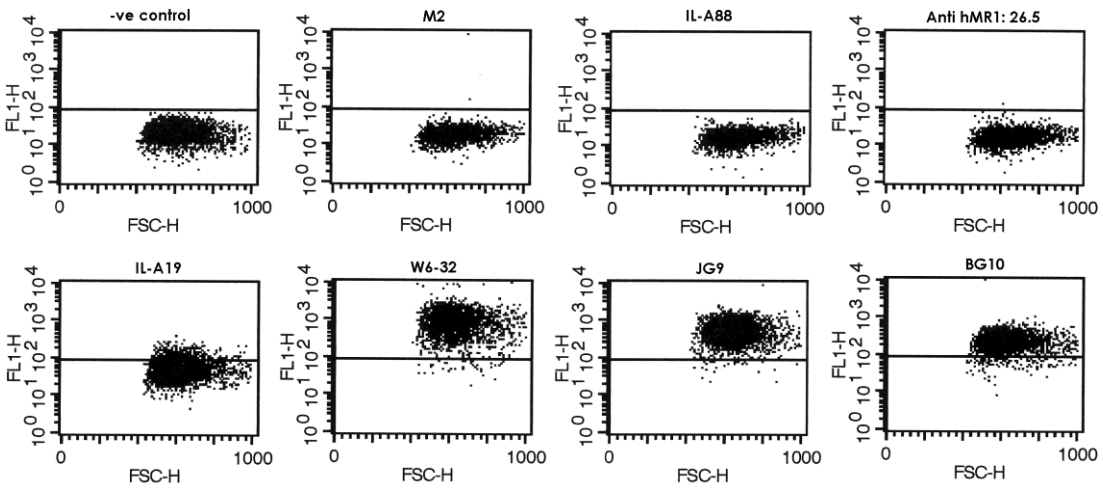


Figure 6-6: Flow cytometry analyses of non-transfected Cos-7 cells stained with antibodies specific for pFlag-CMV-3 vector tag (M2), bovine MHC (IL-A88), human MR1 (26.5), bovine MHC/β2M (IL-A19), human MHC (W6-32), and bovine β2M (JG9 and BG10), and an immunofluorescent secondary antibody. A negative control was incubated with FACS medium and secondary antibody.

Flow cytometry analysis results of transfected Cos-7 cells are summarised in table 6-3. Cells transfected with a classical MHC class I showed a discrete subset of cells

(11%) that stained positively with IL-A88, confirming success of the transfection reaction and cell-surface expression. These cells were negative for 26.5 (see figure 6-7, A). Cos-7 cells transfected with full MR1 did not give any detectable staining with IL-A88, M2 or 26.5 (figure 6-7, B). Cos-7 cells transfected with the MR1/MHC chimaera also did not stain with 26.5 or the M2 antibody, but weak staining of a small percentage of cells (~5%) was observed with IL-A88 (figure 6-7, C).

Table 6-3: Summary of results of % positive fluorescence cells obtained from flow cytometry analyses of untransfected Cos-7 cells and Cos-7 cells transfected with a classical class I MHC, MR1, or an MR1/MHC chimaera.

	Primary Antibodies				
Transfected Construct	**None**	**M2**	**IL-A88**	**IL-A19**	**26.5**
Non-transfected	**0.1%**	**0.1%**	**0.0%**	**16.2%**	**0.1%**
Classical Class I MHC	**0.3%**	**N/A**	**10.7%**	**14.9%**	**0.4%**
Full MR1	**0.3%**	**0.7%**	**0.3%**	**31.7%**	**0.3%**
MR1/MHC Chimaera	**0.7%**	**1.0%**	**5.0%**	**34.0%**	**0.8%**

M2 antibody was specific for the 5' pF3 vector tag; IL-A88 is specific for bovine MHCIa heavy chain; IL-A19 is specific for bovine MHCIa in association with β2M; 26.5 is specific for human MR1 with some reported cross-reactivity between species. The percentage of positively staining cells for each antibody/construct were established by 'gating' cells as positive and negative on mean fluorescence intensity of secondary antibody staining, as shown in figure 6-7.

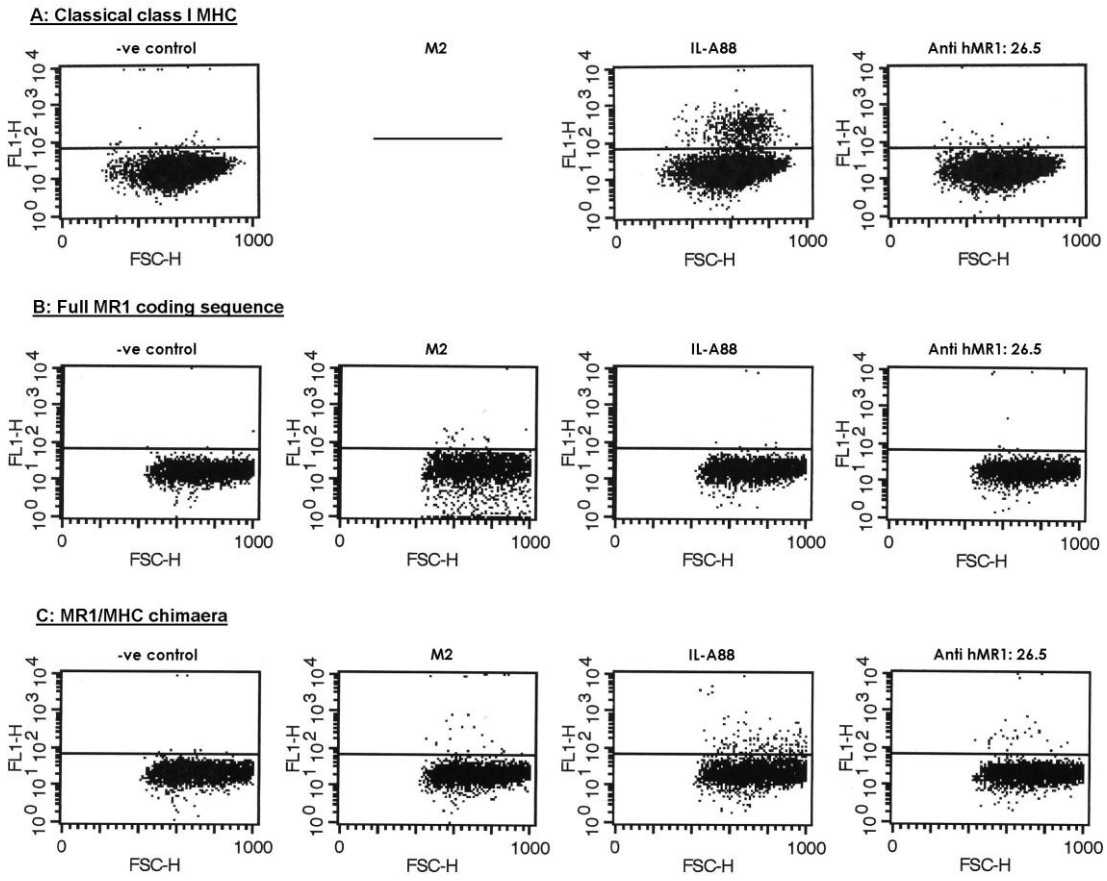


Figure 6-7: Flow cytometry analyses of Cos-7 cells transfected with: A) classical class I MHC in a TOPO TA expression vector, B) full MR1 coding sequence, C) an MR1/MHC chimaera. Primary antibodies used were specific for bovine MHC α3 region (IL-A88) and human MR1 (26.5), as well as pFlag-CMV-3 vector tag (M2); an immunofluorescent secondary antibody was used. A negative control was incubated with FACS medium and secondary antibody.

Results obtained with fixed cells stained by immunofluorescence and analysed by fluorescence microscopy are summarised in table 6-4. Cos-7 cells transfected with a classical class I MHC molecule gave strong staining with IL-A88. Those transfected with full MR1 demonstrated positive fluorescence of cells stained with the tag-specific M2 antibody, but no staining with IL-A88 or the putative MR1-specific antibody 26.5. Cos-7 cells transfected with the MR1/MHC chimaera stained strongly with M2 and also IL-A88 but were still negative for 26.5 (see figure 6-8). These results indicate that IL-A88 is able to detect the MHC α3 region of the chimaera. They also demonstrate that both the MR1 and the MR1/MHC chimaeric constructs

are expressed in transfected cells but fail to detect any reactivity of the 26.5 antibody with bovine MR1.

Table 6-4: Summary of results obtained following immunofluorescence staining of fixed Cos-7 cells transfected with either a classical class I MHC, MR1, or an MR1/MHC chimaera.

	Primary Antibodies				
Transfected Construct	**None**	**M2**	**IL-A88**	**IL-A19**	**26.5**
Classical Class I MHC	**-**	**NT**	**+**	**NT**	**NT**
Full MR1	**-**	**+**	**-**	**-**	**-**
MR1/MHC Chimaera	**-**	**+**	**+**	**-**	**-**

M2 antibody was specific for the 5' pF3 vector tag; IL-A88 is specific for bovine MHCIa heavy chain; IL-A19 is specific for bovine MHCIa in association with β2M; 26.5 is specific for human MR1 with some reported cross-reactivity between species. Positive and negative staining were defined as detectable and non-detectable fluorescence of the secondary antibody, respectively. NT: not tested.

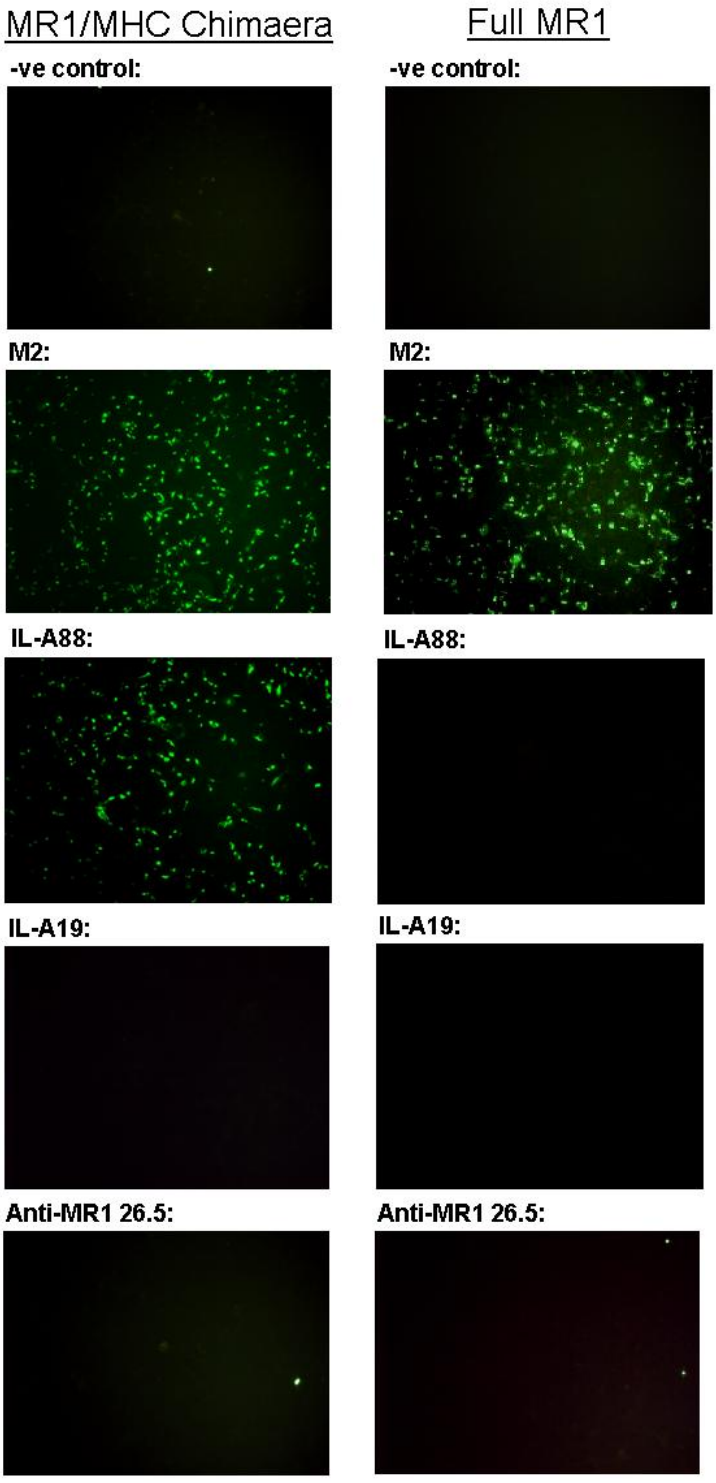


Figure 6-8: Immunofluorescence staining of fixed Cos-7 cells transfected with an MR1/MHC chimaera or full MR1 coding sequence, with primary antibodies specific for the 5' vector tag (M2), bovine MHC heavy chain (IL-A88), human MR1 (26.5), and bovine MHC/β2M complexes (IL-A19). The negative control sample cells were incubated with medium and secondary antibody only.

6.4 Discussion

A number of bovine cell lines belonging to different leucocyte-lineages were tested for expression of MR1 at the mRNA level using RT-PCR techniques. All but one of these lines were transformed by infection with *Theileria* parasites and included lines of the monocyte, B-cell and T-cell lineages. Results showed that, as for humans, MR1 was abundantly transcribed in all cell-lines tested. In addition to the predicted full-length transcript, a second smaller product corresponding in size to the MR1 splice variant lacking α1 exon (identified in chapter 5) was also ubiquitously expressed, although at lower levels. As discussed in the previous chapter, the ubiquitous detection of this alternatively spliced variant suggests that it may have a functional role in regulating expression of the intact protein.

Although MR1 transcripts were readily detected in all of the cell lines tested, they appeared to be more abundant in the two B-cell lines than in the T-cell or monocyte lines. Direct comparison of the B-cell and monocyte lines is particularly valid since these consisted of pairs of lines (B-cell and monocyte) derived from the same animals. Further experiments using quantitative PCR assays applied to a larger number of cell lines are required to quantify these differences in expression. However, given the evidence that in mice expansion of the MAIT cell population is dependent on B-cells (Treiner et al. 2003), these observations suggest that the observed increased transcription of MR1 in B-cell lines may relate to a functional role of MR1 expressed in activated B-cells.

Studies of MR1 expression in ruminants have been limited by lack of specific antibodies. This study tested a number of different antibodies known to react with bovine classical class I MHC, as well as an antibody raised against human MR1, which had been reported, in a paper published during the course of the current studies, to recognise bovine MR1 (Huang et al. 2009). These antibodies were tested for recognition of cells transfected with the full coding sequence of bovine MR1 and with a chimaeric cDNA containing the α3, transmembrane and cytoplasmic segments from a bovine classical class I gene. The latter was included because previous

studies in mice using a similar chimaeric construct (Miley at el, 2003) demonstrated a 3-fold increase in surface expression in transfected cells in comparison to the native MR1. The inclusion of a classical class I α3 domain also provided the possibility that the MR1 chimaera heavy chain might react with the IL-A88 antibody, which is specific for a non-polymorphic determinant on bovine classical class I heavy chains (Toye et al. 1990). The expression vector used in the experiments also encoded an N-terminal peptide tag that would allow detection of both intracellular and cell surface expression. Several of the antibodies examined reacted with the untransfected Cos-7 cells; these included W6-32 which recognises human classical class I and is known to cross-react with class I in some other species (Maziarz et al. 1986). Recognition of the Cos cells by antibody IL-A19, which recognises classical class I heavy chain in association with β2M and has also been shown to recognise other species' heavy chains associated with bovine β2M (Bensaid et al. 1989), as well as the two bovine β2M-specific antibodies, was probably due to association of the monkey class I heavy chains with bovine β2M in the culture medium. Further experiments using serum-free medium or medium supplemented with mouse serum instead of bovine serum are required to confirm this suggestion.

Cells transfected with both MR1 constructs exhibited strong and specific intracellular staining with the antibody directed against the peptide tag, but no cell surface staining was detected with the same antibody. Similar intracellular staining of cells transfected with the chimaeric construct, but not the native MR1, was obtained with antibody IL-A88, confirming that this antibody reacts with the bovine α3 domain of class I heavy chains. Although there was also a suggestion of low level cell surface staining with IL-A88 of a small percentage (~5%) of cells transfected with the chimaeric construct, the absence of staining of these cells with the tag-specific antibody suggested that this may have been non-specific. Hence, although MR1 was clearly expressed intracellularly, no evidence of cell surface expression was detected.

Studies of MR1 expression in human cells and tissues obtained *ex vivo* have yet to provide convincing evidence of cell surface expression of this protein (Treiner et al. 2005). Moreover, *in vitro* transfection and transduction studies with mouse and

human MR1 have reported low and variable levels of cell surface expression, despite readily detectable intracellular protein (Huang et al. 2005). While early studies suggested that MR1 was retained within the endoplasmic reticulum (Miley et al. 2003), a more recent study employing microscopy with markers to sub-cellular compartments has demonstrated that a large quantity of MR1 protein is retained within the late endosomal compartments (Huang et al. 2008). Based on these and other findings, it has been proposed that surface expression of MR1 is strictly regulated, possibly dependent on availability of ligand within an intracellular compartment (see below). The subcellular localisation of MR1 was not examined in the current study.

Classical class I molecules are usually expressed on the cell surface in a folded conformation, which is dependent on the binding of an appropriate peptide ligand. However, under some circumstances they can be expressed in an open conformation in which there is no bound ligand. Previous studies of MR1 demonstrated that human Hela cells and B6/WT-3 cells transduced with MR1 showed low levels of expression of the molecule at the cell surface. By using antibodies that were shown to be specific for open and folded forms of MR1, which are indicative of empty and antigen-bound forms respectively, Huang et al. were able to show that both forms of the protein were present on the cell surface. By using the antibodies to inhibit T cell recognition, they also obtained evidence that only the folded conformer of MR1 was able to activate two MAIT cell hybridomas (Huang et al. 2005). More recently it has been shown that acid eluates from MR1 contain an unidentified self-ligand which enhances activation of MAIT cell hybridomas (Huang et al. 2009). These findings support the hypothesis that MR1 is dependent on endogenous ligand for folding and possibly for efficient cell surface expression.

The monoclonal antibody, 26.5, was raised against a soluble ectodomain of human MR1 (Miley et al. 2003) and has been shown to specifically recognise a fully folded form of the protein, presumed to contain a bound ligand (Huang et al. 2005). Recently, these workers independently reported the isolation of a bovine MR1 cDNA and, based on flow cytometry analyses of human and murine cell lines transduced

with a recombinant retrovirus, they claimed that the 26.5 antibody cross-reacted with bovine MR1 (Huang et al. 2009). The current study failed to detect any reactivity of this antibody against transfected bovine MR1 sequence or the MR1/MHC chimaera, by either intracellular or cell surface staining. This may reflect lack of sensitivity of staining due to lower affinity of the antibody for bovine MR1 compared to human MR1, or it may be due to failure of the antibody to detect intracellular protein that has not adopted a folded conformation. The author's original method of retroviral transduction, rather than transfection, may also have affected the findings.

Hence, the failure to detect cell surface expression in the current study may relate to lack of an appropriate endogenous ligand to allow appropriate folding of the protein and transport to the cell surface. Although it has been hypothesised that microbial colonisation of the gut stimulates cells to produce endogenous MR1 ligands which may allow or increase MR1 cell-surface expression, no experiments on the effects of bacterial extracts on MR1 expression have been performed. Similarly, certain conditions of cellular stress or infection state may be required to stimulate MR1 expression. The ability to express MR1 efficiently on the cell surface may also be dependent on cell type and/or interaction with other cells such as antigen-presenting cells.

7 General Discussion

Mucosal Associated Invariant T-cells are a subset of αβ T-cells characterised by expression of an invariant TCRα chain, and their restriction by the non-classical MHC class I-like molecule, MR1 (Treiner et al. 2003). Both of these elements are highly conserved in humans and mice. The MAIT orthologous TCRα chain rearrangements have a CDR3 of constant length and sequence and are used in combination with a limited repertoire of Vβ chains. MAIT cells are notably more abundant in humans than mice, and are predominantly enriched in the $CD4^{-}CD8^{-}$ T-cell subset of both species (Tilloy et al. 1999). Studies in various gene knock-out mice indicate that MAIT cells undergo stepwise development (Martin et al. 2009); they are initially selected in the thymus, where they are present in small numbers and have a naïve T cell phenotype. Intra-thymic development is dependent on MR1 and a non B-cell, non T-cell, haemopoietic-derived cell-type. They then undergo clonal expansion in the periphery, which is dependent on B-cells and commensal intestinal flora, and adopt a memory phenotype. Mature MAIT cells also display a variety of NK cell surface markers. Following maturation, MAIT cells localise in specific tissues, notably the intestinal mucosa (Treiner et al. 2005). Like cells of the innate immune system, they produce a rapid, diverse cytokine response upon activation (Kawachi et al. 2006). However, the function of MAIT cells remains unclear.

The results obtained in the current study have confirmed the presence of MAIT cells and the restriction element MR1 in ruminants and provided further evidence that this system is highly conserved between mammalian species, suggesting that it plays a key role in the immune system of mammals.

Conservation of the MAIT TCRα and MR1 amino acid sequences: The present study has confirmed the presence of MAIT cells and MR1 in both cattle and sheep. Both the invariant TCRα and MR1 sequences are highly conserved between ruminants and humans and mice, and the TCRα chains use orthologous V and J gene segments. Indeed, the cross-species conservation of MAIT/MR1 sequence is much higher than that seen for other immune elements, including iNKT/CD1d. The CDR

sequences of the MAIT TCRα are also highly conserved, suggesting that they have very similar specificities in the different species. Additionally, the extracellular domains of MR1 (α1, α2 and α3) show a particularly high level of conservation, suggesting that this molecule has a highly conserved function in antigen presentation. Since these domains form the antigen-binding groove of the protein, this suggests that the cognate ligand for MAIT cells is also highly conserved between species. Since murine MAIT cell hybridomas have been shown to respond to both murine and human MR1, the possibility that they also can respond to bovine MR1 could be investigated in future studies. The conservation of MAIT TCR and MR1, coupled with the relatively high frequency of MAIT cells in peripheral lymphoid tissues and evidence from mice that their activation results in a rapid cytokine response, suggests that these cells may be capable of mounting an innate immune response against a conserved ligand. Although the current studies have demonstrated expression of a conserved TCRα chain, the extent to which the associated TCRβ chains are conserved was not investigated. This should be investigated in future work, but will require the development of methods to purify bovine MAIT cells.

Studies in mice have provided evidence that the ligand for MAIT cells is lipid in nature, since they are activated by the synthetic glycolipid αManCer, and αManCer-loaded murine MR1 has been shown to stimulate MAIT cells (Okamoto et al. 2005; Shimamura et al. 2007). Given these findings, and the evidence that MAIT cell peripheral expansion is dependent on commensal bacterial flora, it has been hypothesised that MAIT cells may interact with bacterial cell-wall lipids and play a role in mucosal immune responses and/or prevention of immune responses to commensal gut flora. However, it is interesting to note that although MAIT clonal expansion is dependent on gut flora, none of the published studies of MAIT cell activation have examined responses to bacteria or bacterial products. This provides an area for potential future work on MAIT cells and MR1, which may provide vital clues to the ligand for, and function of, the MR1-MAIT cell system. However, again methods to identify MAIT cells are required to undertake such studies in ruminants.

Expansion and distribution of the bovine MAIT cell population in the post-natal period: The work described in chapter 4 successfully developed quantitative PCRs for measuring MAIT TCRα transcripts and total TCRα transcripts, thus permitting estimation of the percentages of T-cells expressing the MAIT TCR. Using this assay, T-cells in PBMC from adult (>2 years) cattle were estimated to contain 0.1-0.6% MAIT cells. By contrast, very few MAIT cells were detected in PBMC and tissues of neonatal calves. However, MAIT cell levels were already significantly increased by 3-weeks of age and the levels were maintained in 3-month old animals. Despite this increase, the levels of MAIT TCR in thymus remained low in all age groups. These results support the evidence in humans and mice that MAIT cells undergo peripheral expansion, and suggest that this occurs in the immediate post-natal period. Consistent with findings in humans and mice (Tilloy et al. 1999; Treiner et al. 2003; Treiner et al. 2005), relatively high levels of MAIT were found in some gut associated lymphoid tissues. However, the highest levels of MAIT cells (>2%) were found in the spleen of post-neonatal calves, and the spleen was also a rich source of MAIT cells in sheep. This raises the interesting question of where the peripheral expansion of MAIT cells occurs. Since this process has been shown in mice to be dependent not only on MR1 but also on the presence of B cells and an intact gut flora (Tilloy et al. 1999; Treiner et al. 2003), it has been hypothesised that expansion may occur in the intestinal mucosa. However, it is possible that commensal flora may serve to induce MR1 expression on B-cells (through either a bacterial ligand or stimulation of an endogenous ligand) and that these MR1-expressing B cells migrate to other lymphoid tissues including the spleen, which is a highly vascular organ and contains numerous B-cell follicles (van Ewijk et al. 1977; Nieuwenhuis and Opstelten 1984). Thus, expansion of the MAIT cell population could occur throughout the lymphoid system but be most pronounced in tissues, such as the spleen, that are rich in activated B cells.

It would be of interest to determine whether MAIT cells are also abundant in human and mouse spleen. However, because mice are kept in very clean environments, the spleens of healthy animals tend to contain very few germinal centres or activated B cells. Thus, if the presence of activated B cells is important for expansion or accumulation of MAIT cells, the findings may not be the same in the mouse. Indeed,

it is possible that the overall lower numbers of MAIT cells seen in mice is related to a less active immune system than that in outbred species. This could be further assessed in mice by examining the numbers of MAIT cells in animals experimentally immunised to activate the immune system.

Bovine MAIT cells are enriched in the CD8$^+$ T-cell population. The small CD4$^-$ CD8$^-$ T-cell populations in humans and mice are both enriched for MAIT cells in comparison to other T-cell populations, although significant populations are also found in the CD8αβ$^+$ T-cell population in humans, and CD4$^+$ T-cells in mice (Tilloy et al. 1999; Martin et al. 2009). However, neither qPCR nor sequencing of multiple cDNAs revealed any evidence of enrichment of MAIT TCRα expression in this small population of bovine PBMC. Instead, bovine MAIT TCRα transcripts appeared to be enriched in the CD8$^+$ T-cell population. These results suggest that bovine MAIT cells may utilise CD8 as a co-receptor, in which case bovine CD8 would be predicted to interact with MR1. However, no evidence has been found for the interaction of human (or murine) CD8 with MR1, and the residues at positions 223 to 229, 245, and 247 of the classical class I MHC α3 domain, which are indispensable for interaction with the CD8 coreceptor (Sun et al. 1995; Neveu et al. 2006) and are present in both human and murine MHC-encoded class I molecules, are absent from human and murine MR1 (Riegert et al. 1998). Comparison of bovine MR1 α3 region and bovine classical class I MHC α3 sequences shows similar disparity in these residues; only 2 of the 9 residues are conserved, suggesting that MR1 is unlikely to interact with CD8.

The CD8 protein comprises α and β chains and most T-cells express both chains, which combine to produce a surface heterodimer. However, the α chain can also generate a homodimeric form of the protein and a small subset of T-cells have been shown to express the CD8α but not the CD8β chain (Wheeler et al. 1992). In cattle such cells are identifiable using the α chain-specific monoclonal antibody CC-58 (MacHugh and Sopp 1991). Further studies are required to determine whether the bovine MAIT cells express CD8α only or CD8αβ. Provision of more information on the phenotype and functional characteristics of MAIT cells is dependent on the

development of a monoclonal antibody specific for the MAIT TCRα chain, specifically the MAIT CDR3 region. This could be attempted using either recombinant bovine MAIT TCRα chain protein (or the extracellular domains) or synthetic peptides representing the CDR3 of the α chain as immunogens in mice. Specific antibodies would be expected to detect a small but discrete population of T-cells in PBMC or lymphoid tissues.

Detection of abundant alternatively spliced transcripts of bovine MR1. The consistent detection of MR1 transcripts in PBMC and in each of a number of different bovine cell lines tested suggests that, as in other species (Hashimoto et al. 1995; Yamaguchi et al. 1997; Riegert et al. 1998), the MR1 gene is ubiquitously transcribed. In addition to the full-length transcript, shorter alternatively spliced forms, lacking either the α1 exon, or α1 and α3 exons, were consistently detected. These bore similarity to alternatively spliced forms of MR1 detected in humans and mice, in which all or part of the α1 or α3 exons are omitted (Riegert et al. 1998). All detected ruminant splice variants retained an intact open reading frame and would be expected to generate a polypeptide. The ubiquitous transcription and abundance of the splice variant lacking the α1 exon suggests that this molecule may have some functional role. Although it is unlikely to be able to present antigen, it may somehow play a role in regulating expression of the intact MR1 on the cell surface.

Further work is required to investigate the functional significance of these alternatively-spliced MR1 transcripts. Quantitative PCR could be employed to quantify expression in cells of different lineages and to examine the effects of different activation stimuli on the levels of expression compared to that of the intact MR1 transcript. Also, if an in vitro expression system were to be developed for bovine MR1, it would be possible to examine the effect of co-transfection with the alternatively-spliced cDNAs on expression of MR1 protein.

Transfected MR1 is expressed intracellularly but not on the cell surface. Results in chapter 6 demonstrated successful transfection of bovine MR1 in Cos-7 cells; however, transfected MR1 and an MR1/MHC-chimaera were only expressed

intracellularly. The latter was demonstrated using antibody to an N-terminal peptide tag and also using a monoclonal antibody that recognised the MHC α3 domain in the chimaeric protein. Previous transfection studies with mouse MR1 had demonstrated low levels of cell surface expression (<10% of transfected cells) in some cell lines, and that expression was increased to 26% following transfection of an MR1/MHC chimaera containing the α3, transmembrane and cytosolic domains of a classical class I MHC protein (Miley et al. 2003). Later studies showed that the majority of the MR1 protein in transfected cells was retained intracellularly in the late endosomal compartment (Huang et al. 2008). The work in chapter 6 did not include analysis of the subcellular site of MR1 protein expression. Further work applying two-colour confocal microscopy using markers for sub-cellular compartments is required to define the intracellular sites of MR1 protein localisation. Transfection of different cell lines should also be investigated, to determine whether cell type and species of transfected cells influences expression. Given the evidence that B-cells influence peripheral expansion of MAIT cells (Treiner et al. 2003), these experiments should include transfection of bovine B-cell lines. The effects of various cell activation stimuli on expression could also be examined. Success in any of these reactions would enable experiments to define precisely what determines cell surface expression, and would provide cells that could be used to test the responsiveness of MAIT cells and ultimately analysis of potential ligands.

Such studies of MR1 would be aided by the development of a specific monoclonal antibody. The studies with cells transfected with bovine MR1 failed to confirm that the human MR1-specific monoclonal antibody 26.5 cross-reacts with the bovine protein. Production of a specific antibody could use recombinant polypeptide incorporating the extracellular domains of MR1 as an immunogen, as has been used successfully for human MR1 (Huang et al. 2005). A monoclonal antibody specific for MR1 could be used to answer a number of unresolved questions: Which cell types in the thymus express MR1 during selection of MAIT cells? Can cell surface expression of MR1 be detected in vivo and on what cell types? What factors regulate cell surface expression and do the transcriptional splice-variants of MR1 play a role with regard to function or regulation of expression of the intact protein?

In summary, ruminant MAIT cells and MR1 appear to share many features with their human and murine orthologues. The cross-species conservation of both is startling, and suggests evolutionary conservation of structure, and therefore, potentially, function. The ruminant MR1-MAIT cell system is more similar to the human system than mice, where MAIT cells remain in low numbers in the periphery, and may therefore provide a better functional model to identify the ligands and functions of these molecules/cells.

Post-script: Following submission of this thesis, two publications (Gold et al. 2010; Le Bourhis et al. 2010) have demonstrated *in-vitro* activation of human and mouse MAIT-cells by antigen-presenting cells infected with various strains of bacteria and yeast, but not virus, in an MR1-dependent manner. Activation was independent of both classical MHC class I and II pathways, as well as the main innate immune receptors. Additionally, numbers of circulating MAIT cells were reduced in the blood of patients with certain bacterial infections, corresponding with migration of these cells into infected tissues. These results suggest that MAIT cells are involved in either the direct control, and/or the subsequent development of adaptive immunity, to bacterial infection.

8 Appendix A: Materials and Methods

8.1 Solutions and media

8.1.1 Alsever's Solution 10x stock

D-glucose	205.0g	(113.8mM)
Citric acid	5.5g	(2.9mM)
Sodium chloride	42.0g	(71.9mM)
Tris-sodium citrate di-hydrate	80g	(27.2mM)

Made up to 1 litre with double-distilled water (DDW).

Diluted to 1x solution by adding 100ml stock solution (filter sterilised through 0.45µm Minisart single-use filter (Sartorius, Goettingen, Germany)) to 900ml sterile DDW.

8.1.2 RBC Lysis Buffer

Buffer composed of 1 part 0.175M TRIS pH7.4 and 9 parts 0.16M ammonium chloride, and filter sterilised through 0.45µm Minisart single-use filter.

Trizma base	10.6g	(0.175M)

pH adjusted by addition of concentrated HCl to pH7.4; made up to 500ml volume with DDW.

Ammonium chloride	8.55g	(0.16M)

Made up to 1 litre volume with DDW.

8.1.3 Standard Culture Medium

RPMI 1640 medium + 25mM HEPES + L-Glutamine (Gibco, Paisley, UK)

10% foetal bovine serum (Gibco, Paisley, UK)

100u/ml penicillin, 100µg/ml streptomycin, 292µg/ml L-glutamine, as Penicillin-Streptomycin-Glutamine (100x) solution (Gibco, Paisley, UK)

2-Mercaptoethanol (2ME) 5 x 10^{-5}M

8.1.4 DMEM-Cos Medium

DMEM (Gibco, Paisley, UK)

(Dulbecco's Modified Eagle Medium 1X,

+ 4.5g/L Glucose + L-Glutamine + Pyruvate)

10% foetal bovine serum (Gibco, Paisley, UK)

100u/ml penicillin, 100µg/ml streptomycin, 292µg/ml L-glutamine, as Penicillin-Streptomycin-Glutamine (100x) solution (Gibco, Paisley, UK)

2-Mercaptoethanol (2ME) 5 x 10^{-5}M

8.1.5 Phosphate Buffered Saline (PBS) / Ethylenediaminetetraacetic acid (EDTA)

25ml 20 x PBS stock solution and 54ml 0.5M EDTA stock solution mixed into 400ml DDW. Adjusted to pH7.2; made up to 500ml volume with DDW; filter sterilised through 0.45µm Minisart single-use filter. Stored at 4°C. PBS tablets used in accordance with manufacturer's instruction (Unipath Ltd., Bedford, UK)

0.5M EDTA: Na2 EDTA 2 H20 93.05g (0.5M)

Soluted in 400ml DDW; adjusted to pH8 with sodium hydroxide; made up to final volume 500ml with DDW; filter sterilised through 0.45µm Minisart single-use filter.

8.1.6 FACS Medium

RPMI 1640 medium + 25mM HEPES + L-Glutamine (Gibco, Paisley, UK)

2% foetal bovine serum (Gibco, Paisley, UK)

0.2% sodium azide

8.1.7 Phosphate Buffered Saline

PBS tablets used in accordance with manufacturer's instruction (Unipath Ltd., Bedford, UK)

8.1.8 SM-0005 PCR Buffer

Produced by Abgene, Surrey, Uk.

45mM Tris-HCl (pH8.8 at 25°C)

11mM ammonium sulphate

4.5mM magnesium chloride

0.113mg/ml bovine serum albumin

4.4µM EDTA

1.0mM each of dATP, dCTP, dGTP, dTTP

8.1.9 Tris-acetate/EDTA (TAE) electrophoresis buffer (50x stock)

Trizma base	242g	(2M)
Glacial acetic acid	57.1ml	(2M)
EDTA	100ml of 0.5M solution	(50mM)

Adjusted to pH7.0 with 5M sodium hydroxide; made up to 1 litre volume with DDW; autoclaved.

8.1.10 Loading Buffer for Agarose Gel Electrophoresis

15% ficoll (Type 400) in DDW

0.25% bromophenol blue

0.25% xylene cyanol

8.1.11 20% Polyethylene glycol (PEG) / 2.5M NaCl

Sodium chloride	58.44g	(2.5M)
PEG (MW8000)	100g	(20%)

Made up to 0.5 litres volume with DDW.

8.1.12 5 x CSA Buffer

Trizma base	4.840g	(400mM)
Magnesium chloride	0.203g	(10mM)

Adjusted to pH9 with concentrated HCl; made up to 100ml volume with DDW; autoclaved.

8.1.13 3M Sodium acetate pH5.2

Sodium acetate.$3H_20$	40.81g	(3M)

Adjusted to pH5.2 with glacial acetic acid; made up to 100ml volume with DDW; autoclaved.

8.1.14 SOC Medium

Tryptone	20g
Yeast extract	5g
Sodium chloride	0.5g
Potassium chloride	0.19g
Glucose	4g

Adjusted to pH7.0 with 5M sodium hydroxide; made up to 1 litre volume with DDW; autoclaved.

8.1.15 LB agar plates with ampicillin

Sodium chloride	10g
Yeast extract	5g
Tryptone	10g
Agar	7.5g

Adjusted to pH7.0 with 5M sodium hydroxide; made up to 1 litre volume with DDW; autoclaved. Melted for use; once solution temperature <55°C, ampicillin was added to give a final concentration of 100µg/ml. Melted media was poured into sterile petri dishes to set.

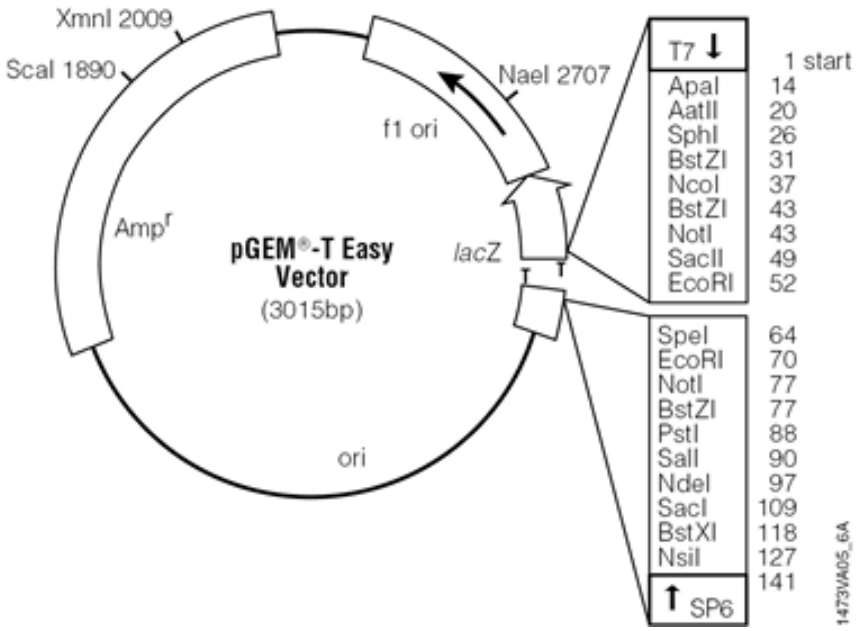
8.1.16 LB medium with ampicillin

Sodium chloride	10g
Yeast extract	5g
Tryptone	10g

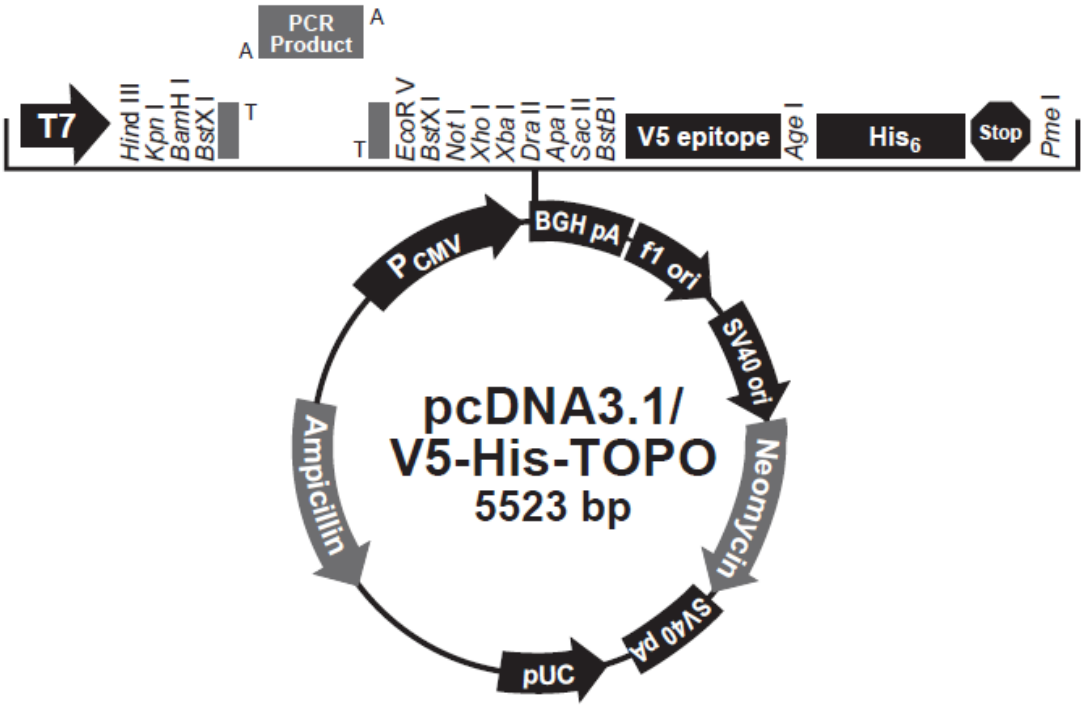
Adjusted to pH7.0 with 5M sodium hydroxide; made up to 1 litre volume with DDW; autoclaved. Ampicillin was added before use to give a final concentration of 100µg/ml.

8.2 Vector Maps

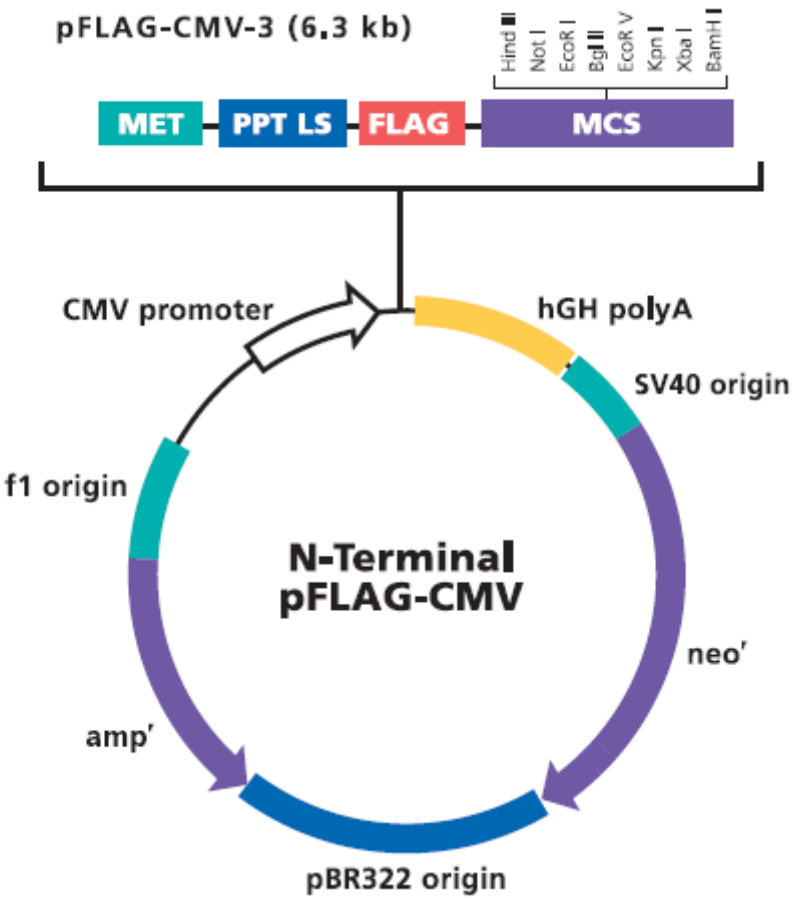
8.2.1 pGEM-T Easy vector



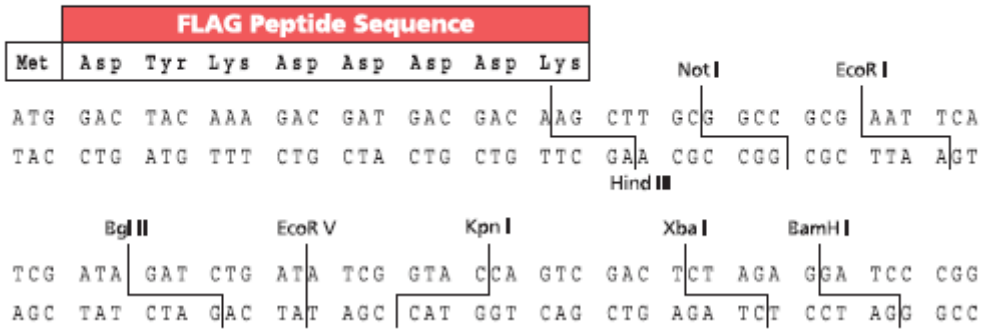
8.2.2 pcDNA3.1 / V5-His TOPO vector



8.2.3 pFlag-CMV-3 expression vector



Multiple Cloning Site
(pFLAG-CMV-3)



9 Appendix B: TCR Vα nomenclature

The table below shows the corresponding names between human TCR Vα genes according to the IMGT and WHO-IUIS nomenclature systems, and is derived from the IMGT website (http://imgt.cines.fr). Genes are presented in the 3' to 5' order in which they are arranged in the human TCRα locus.

IMGT (TRAV)	**WHO-IUIS (TCRVα)**
AV41	19S1
AV40	31S1
AV39	27S1
AV38-2/DV8	14S1-ADV14S1
AV38-1	14S2
AV37	-
AV36/DV7	28S1-DV28S1
AV35	25S1
AV34	26S1
AV26-2	4S1
AV33	-
AV32	-
AV31	-
AV30	29S1
AV29/DV5	21S1-ADV21S1
AV28	-
AV27	10S1
AV8-7	-
AV26-1	4S2
AV25	32S1
AV24	18S1
AV23/DV6	17S1-ADV17S1
AV22	13S1
AV21	23S1
AV20	30S1
AV19	12S1
AV18	-
AV17	3S1
AV16	9S1
AV8-6	1S3
AV12-3	2S2
AV15	-
AV9-2	22S1
AV14/DV4	6S1-ADV6S1
AV13-2	8S2
AV8-5	-
AV8-4	1S2
AV12-2	2S1
AV13-1	8S1
AV8-3	1S4
AV8-2	1S5
AV12-1	2S3
AV11	-
AV10	24S1
AV9-1	-
AV8-1	1S1
AV7	-
AV6	5S1
AV5	15S1
AV4	20S1
AV3	16S1
AV2	11S1
AV1-2	7S2
AV1-1	7S1

10 Appendix C: Cross-species comparison of MAIT TCRα nucleotide sequences

Nucleotide sequences for the full coding regions of the MAIT TCRα of sheep, cattle, humans and mice are shown on the following page. A"." is used to indicate when a sequence for comparison contains the exact same base pair at that location as the top sequence.

```
                                10                  20                  30                  40                  50                  60
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      G C T C T G C A G G  A A A A G G C G T T  G A G C A G C C C A  C T G A G T T G A C  G G T T A T G G A A  G G A G C C T C T G
Bovine MAIT     . . . . . .  . . . . . .  A . . . . . .  . . . . . .  . . T  . . C . . . A . . .  . . . . . .
Human MAIT      . . A . . A . . . . .  . C . . A A . A . .  . . . C . . . . .  . . . . . A . .  A . C . . C . . . .  . . T . . . A T . .
Murine MAIT     . T G . . . . . .  . C . G . . T . . G  . . . . . .  . . T G  . C A . A . . . .  T  . T C . G . . . .  G  . . . A . . . T . .

                                70                  80                  90                  100                 110                 120
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      C C C A G G T C A A  C T G C A C C T A C  C A G A C A T C T G  G G T T C A A T G G  A C T G T T C T G G  T A C C A G C G A C
Bovine MAIT     . . . . . .  T . .  . . . . . .  . . . . . .  . . . . . .  . A . . . . . .  . . . . . .  . . . . . .  . . . . . .  A . .
Human MAIT      T . . . . A . . . .  . . . . . .  . . . . . G . . .  . . . . . .  . . . . . .  C . .  G . . . . . .  . . . . . .  A . .
Murine MAIT     . T . G . . . . . .  . . . . . .  A . . .  A G C . . C . . A .  . . . . . .  . . . .  C . .  G T . A . C . . . .  . . . . . .  A . .

                                130                 140                 150                 160                 170                 180
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      A T G A T G G T G G  A G C A C C T G T G  T T T C T C T C T T  A C A A T G T T T T  G G A T G G T T T G  G A G A C G A G A G
Bovine MAIT     . C . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . C .  . . . . . .  . . . . . .  . . . . . .
Human MAIT      . . . C . . . . C . A  . . . . . .  . . . . . G . . . .  . . . . . .  . . . . . C .  . . . . . .  . . . G A . . A . .
Murine MAIT     G . . . A . . C C A  . . . C . . . . . A  . . . . . .  . . . . . T .  . T G T . . . . .  . . . . . .  A . . G A C . . T .

                                190                 200                 210                 220                 230                 240
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      G T C A T T T T T C  T T C A T T C C T T  A G A C G C T C T G  A T G C A C A C A G  T T A C C T C C T T  C T G A A G G A A C
Bovine MAIT     . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .
Human MAIT      . . . G . . . . . .  . . . . . .  . . T . . G . . . .  A . G G T . . . . .  . . . . . .  T . . . . . .  . . . . . G .
Murine MAIT     . G . . . . . C . .  C A . T . . . . . G  . . C . . . . . .  G A  . . . G G T . . . . .  . . . . . .  . . . . C A . . G .

                                250                 260                 270                 280                 290                 300
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      T C C A C A T G A A  A G A C T T T G C C  T C T T A C C T C T  G T G C T G T G A T  G G A T G G C A A C  T A T C G G T T G A
Bovine MAIT     . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . T . . . . . .  . . . . . .  . . . . A . . G . .
Human MAIT      . . . . G . . . . . .  . . . . . C . . . .  . . . . . .  . . . . . .  . . . . . A  . . . . A . . . . . .  A . . . A . .
Murine MAIT     . . . . G . . C . . .  . . . . . C . . . .  . . A . . . . . .  . . . . . .  . . . . . G  . . . . A . . . . . .  A . . . . .

                                310                 320                 330                 340                 350                 360
                     .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Ovine MAIT      T C T G G G G C T C  T G G G A C C A A G  C T A A T T A T A A  A G C C A G
Bovine MAIT     . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .
Human MAIT      . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .
Murine MAIT     . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .  . . . . . .
```

11 Appendix D: Cross-species comparison of MR1 nucleotide sequences

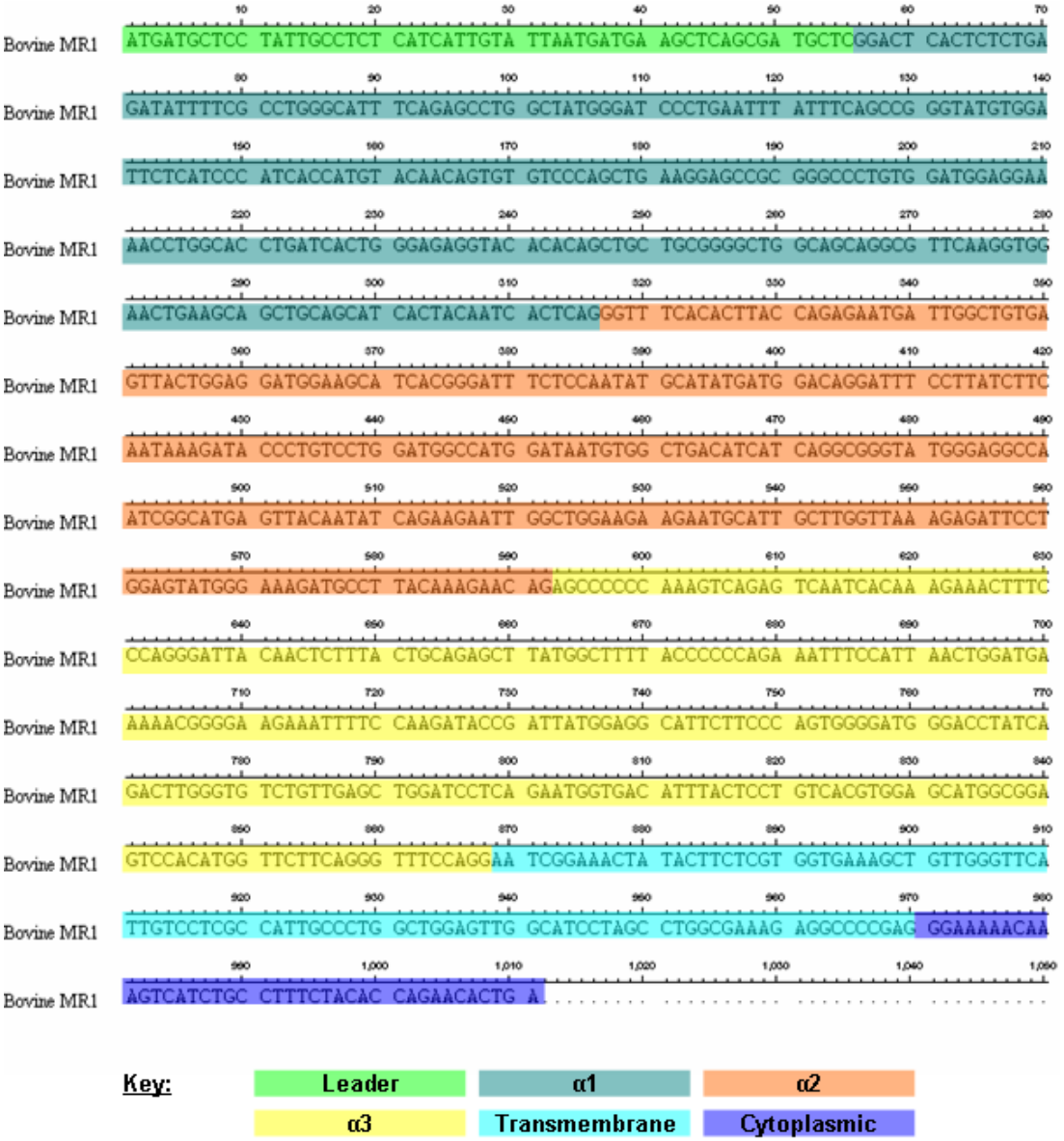
Nucleotide sequences for the full coding regions of MR1 of sheep, cattle, humans and mice are shown on the following two pages. A"." is used to indicate when a sequence for comparison contains the exact same base pair at that location as the top sequence, whereas a "-" indicates a gap in that sequence at that location.

```
                    10        20        30        40        50        60        70        80
                    |         |         |         |         |         |         |         |
Ovine MR1   ATGATGCTCC TATTGCCTCT CATCATTGTA TTAATGATGA AGCTCAGCGA TGCTCGGACT CACTCTCTGA GATATTTTCG
Bovine MR1  ..........  ..........  ..........  ..........  ..........  ..........  ..........  ..........
Human MR1   ...GC.T...  .G..A.....  ..........  .....G....  ...A......  .T.C......  G.........  ..........
Murine MR1  ..........  .G..A.....  GC..GC....  ..CT..G...  ...GA...C.  .A........  C.....G...  ..........

                    90        100       110       120       130       140       150       160
                    |         |         |         |         |         |         |         |
Ovine MR1   CCTGGGCATC TCAGAGCCTG GCTATGGGAT CCCTGAATTT ATTTCAGCTG GGTATGTGGA TTCTCATCCC ATCACCATGT
Bovine MR1  .........T  ..........  ..........  ..........  .......C..  ..........  ..........  ..........
Human MR1   .....G.T..  .G..T..CA  T.C.....G.  ..........  .....G.T..  ....C.....  C..G...C..T  .........CA.
Murine MR1  T....CTG.T  ..C..T....  .TCCC.TCG.  ..........  ..C..T.T..  ..........  C..A...C..T  .....T.CA.

                    170       180       190       200       210       220       230       240
                    |         |         |         |         |         |         |         |
Ovine MR1   ACAACAGTGT CTCCCAGCTG AAGGAGCCAC GTGCCCCATG GATGGCGGAA AACCTGGAAC CTGATCACTG GGAGAGGTAC
Bovine MR1  ..........  G.........  .....G...  .G.....TG..  ......A....  .......C..  ..........  ..........
Human MR1   .TG.......  .A.T.G..A.  ..........  .G........  ......A..G  .....C.CG.  ..........  ..........
Murine MR1  ..G.......  .A.T.GA.A.  ..........  AA..T.....  ......A..G  ........C..  ..........  ..........

                    250       260       270       280       290       300       310       320
                    |         |         |         |         |         |         |         |
Ovine MR1   ACACAGCTGC TGCGGGGCTG GCAGCAGGCG TTCAAGGTGG AACTGAAGCA GCTGCAGCAT CACTACAATC ACTCAGGGTT
Bovine MR1  ..........  ..........  ..........  ..........  ..........  ..........  ..........  ..........
Human MR1   .T........  ..A.......  .....AT...  ..........  .........G  C..A...AGG  ..........  .........C
Murine MR1  .T........  .AA.......  .....A.A..  .....C....  .G....G...  C..A...AGG  .....C....  ..........

                    330       340       350       360       370       380       390       400
                    |         |         |         |         |         |         |         |
Ovine MR1   TAACACTTAC CAGAGAATGA TTGGCTGTGA GTTACTGGAT GATGGCAGCA CCACGGGATT TCTCCAATAT GCGTATGATG
Bovine MR1  .C........  ..........  ..........  .........G  .....A....  T.........  ..........  ..A.......
Human MR1   .C........  ..........  ..........  .C.G......  .....A....  ....A.....  ....G..G..  ..A.....C.
Murine MR1  .C....C...  ..........  ..........  ...G..A..A  ..........  ....A..G..  .....G...  ..A.....

                    410       420       430       440       450       460       470       480
                    |         |         |         |         |         |         |         |
Ovine MR1   GACAGGATTT CATTATCTTC AATAAAGATA CCCTCTCCTG GATAGCCATG GATAATGTGG CTAACATCAT CAGGCGGGCA
Bovine MR1  .C........  ..........  ..........  .....G....  ...G......  ..........  ..G.......  .........T.
Human MR1   .C.G......  ..........  .....C....  ..........  .C.G..TG.A  ..........  ..C....C...  ..A..A....
Murine MR1  ...C......  ..........  .....C....  ..........  .T.G......  ...T......  ..C.......  C..A..AA...

                    490       500       510       520       530       540       550       560
                    |         |         |         |         |         |         |         |
Ovine MR1   TGGGAGGCCA ATCGGCATGA GTTACAATAT CAGAAGAATT GGCTGGAAGA AGAATGTATT GCTTGGTTAA AGAGATTCCT
Bovine MR1  ..........  ..........  ..........  ..........  ..........  .....C....  ..........  ..........
Human MR1   ..........  ...A......  ...G.TG...  ..A.......  ..........  ..........  ..C...C...  ..........
Murine MR1  ..........  .C.T......  ...G.....C  ..A.....C.  ..........  .....C....  ..C...C...  ....G...T.
```

```
                    570       580       590       600       610       620       630       640
                    |         |         |         |         |         |         |         |
Ovine MR1   GGAGTATGGG AAAGATACCT TACAAAGGAC AGAGCCCCCC AAAGTCCGAG TCAATTACAA AGAAACTTTC CCAGGGATTA
Bovine MR1  .......... .......G.. .......A.. .......... .....A.... .....C.... .......... ..........
Human MR1   .......... .......C.. .......A.. ..........A CTG...A.... .A...CG.... ..........T .......G..
Murine MR1  ...A...... .G.....C.. ..G.....A.. ......AT..A GTG..AA..A CA.C.CGA.. G..........T ..........

                    650       660       670       680       690       700       710       720
                    |         |         |         |         |         |         |         |
Ovine MR1   CAACTCTTTA CTGCAGAGCT CATGGCTTTT ACCCCCCAGA AATTTCCATA AACTGGATGA AAAACGGGGA AGAAGTTGTC
Bovine MR1  .......... .......... T......... .......... .......T.. .......... .......... ....A..T..
Human MR1   .G.....C.T .......A.. .......... .......... .......A... .G.CA..... .......... ....A.....
Murine MR1  .....T.C.T .......... .......C.. ....A..... .......... .G.CA..... ....T..... ....A...C.

                    730       740       750       760       770       780       790       800
                    |         |         |         |         |         |         |         |
Ovine MR1   CAAGACACCA ATTATGGAGG CATTCTTCCC AGTGGGGATG GGACCTATCA GACTTGGGTG TCTGTTGAGC TGGATTCTCA
Bovine MR1  .....T....G .......... .......... .......... .......... .......... .......... .....C....
Human MR1   .....A.TTG .......... .......... .......... .A.......... .G.G....CA ..AA..... .T...C....
Murine MR1  .....AGTGG .....C..... GG.A..... ..C..... .A.......... A..G...C.. ..A...A.T. .....C....

                    810       820       830       840       850       860       870       880
                    |         |         |         |         |         |         |         |
Ovine MR1   GAACGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGC GTCCACATGG TTCTTCCGGG TTTCCAGGAT TCGGAAACTA
Bovine MR1  ...T..... .......... .......... .......... .......... .......A... .......... ..........
Human MR1   .G.A.CA.. C.......... .....T..... ...CT....T .......... ......A..T CCC.....A ..A.......
Murine MR1  .G.AA....T G.....T..T. .....T..... ...CT.T..T CG...A.... ....GGA..C CCCT.G...A ..A.GGGAC.

                    890       900       910       920       930       940       950       960
                    |         |         |         |         |         |         |         |
Ovine MR1   TGCTTCTGGT GATGAAAGCT GTT--GGGT TCATTGTCCT CGCCATTGCC CTGGCTGGAG TTGGCATCTT AGCCTGGCGA
Bovine MR1  .A.....C.. .G........ ...---.... .......... .......... .......... .....C.... ..........
Human MR1   .C.C....T.. .......... ..CTCT.... C.......... T.T......TG .......... ....TG.TC. ..T....A..
Murine MR1  .C.....GA.. ..GC.CGATC ---TCT..AA C..CAA.T.. .AT....C... .......... ....TG.TC. GAT....A..

                    970       980       990       1000      1010      1020      1030      1040
                    |         |         |         |         |         |         |         |
Ovine MR1   AGGAAGCCCC GAGGGGAAAA CAAAGTCATC TACCTTTCTA CACCAGAACA CTGA
Bovine MR1  .A..G..... .....A.... .......... .G.......... .......... ....
Human MR1   ..A.G..... ...A.C.... TGG...C.... .......C.A. .........T.G A...
Murine MR1  ....-T.A. A..AACT.... AG.......G ....AAC.C. .....AGTG. A...
```

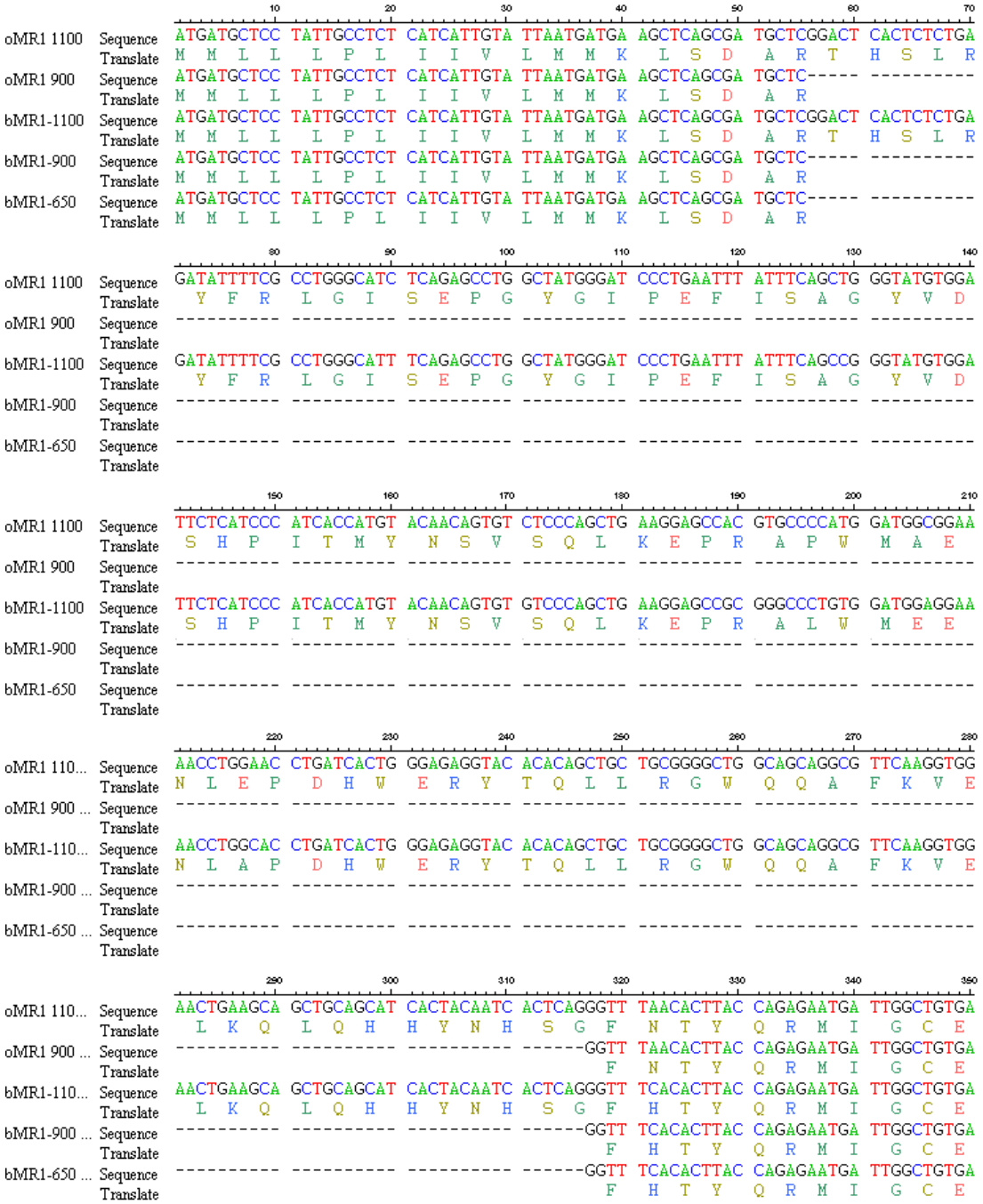
12 Appendix E: Exons of bovine MR1

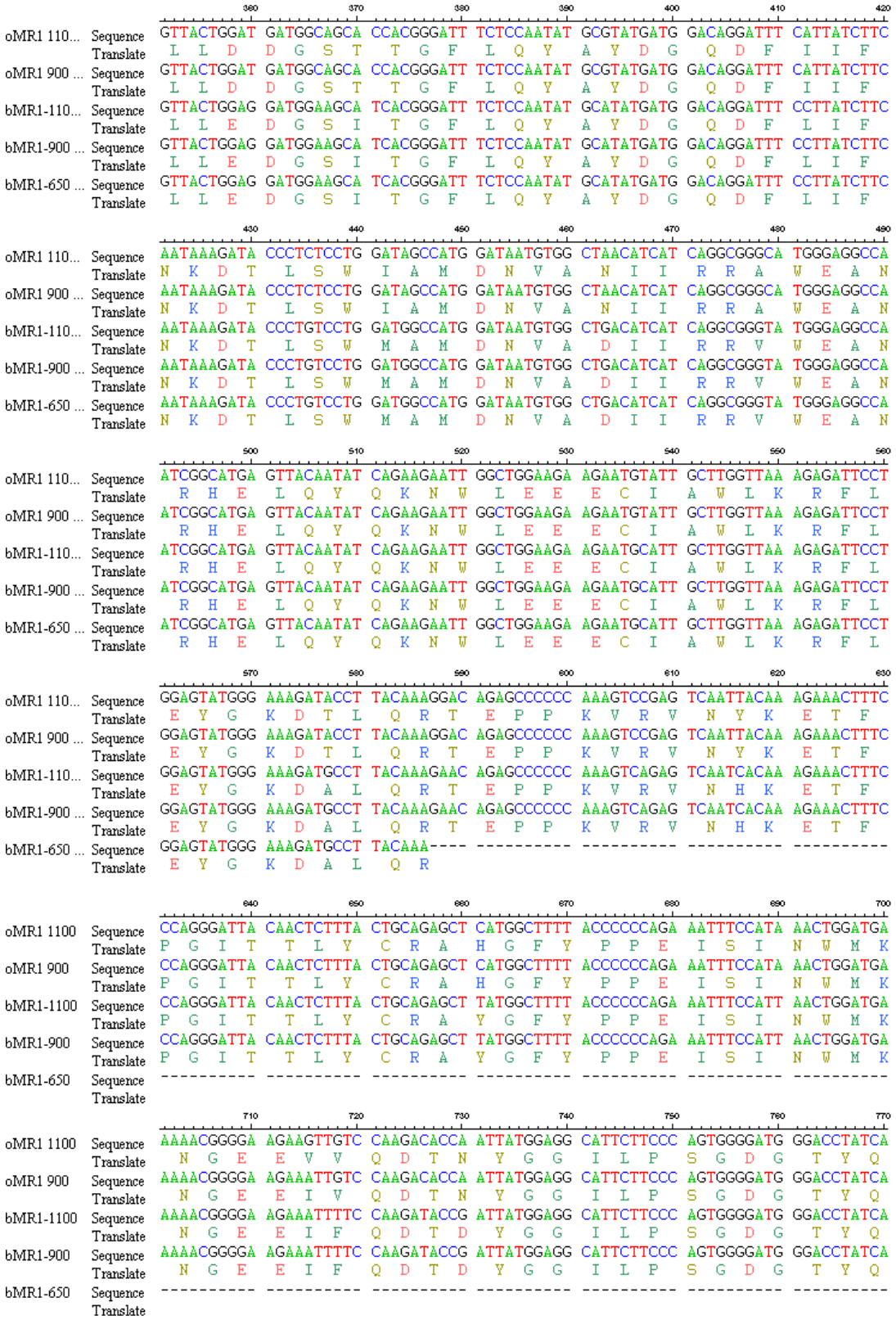
Nucleotide sequences for the exons within the full coding sequence of bovine MR1 are shown below:

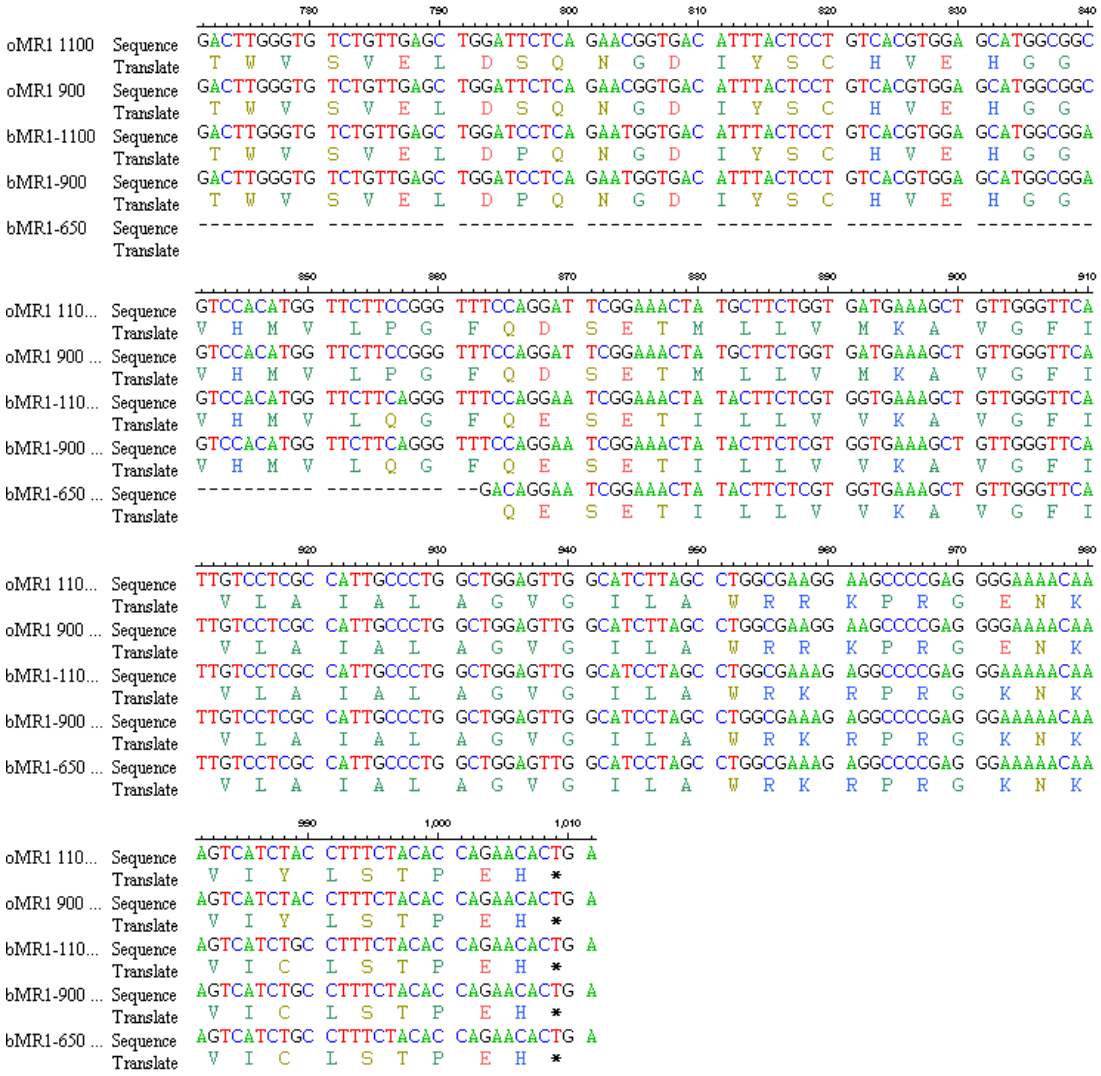


13 Appendix F: Splice variants of bovine and ovine MR1

Nucleotide sequences for coding sequences of the identified splice variants of bovine and ovine MR1 are shown below, in comparison to full MR1 coding sequences:

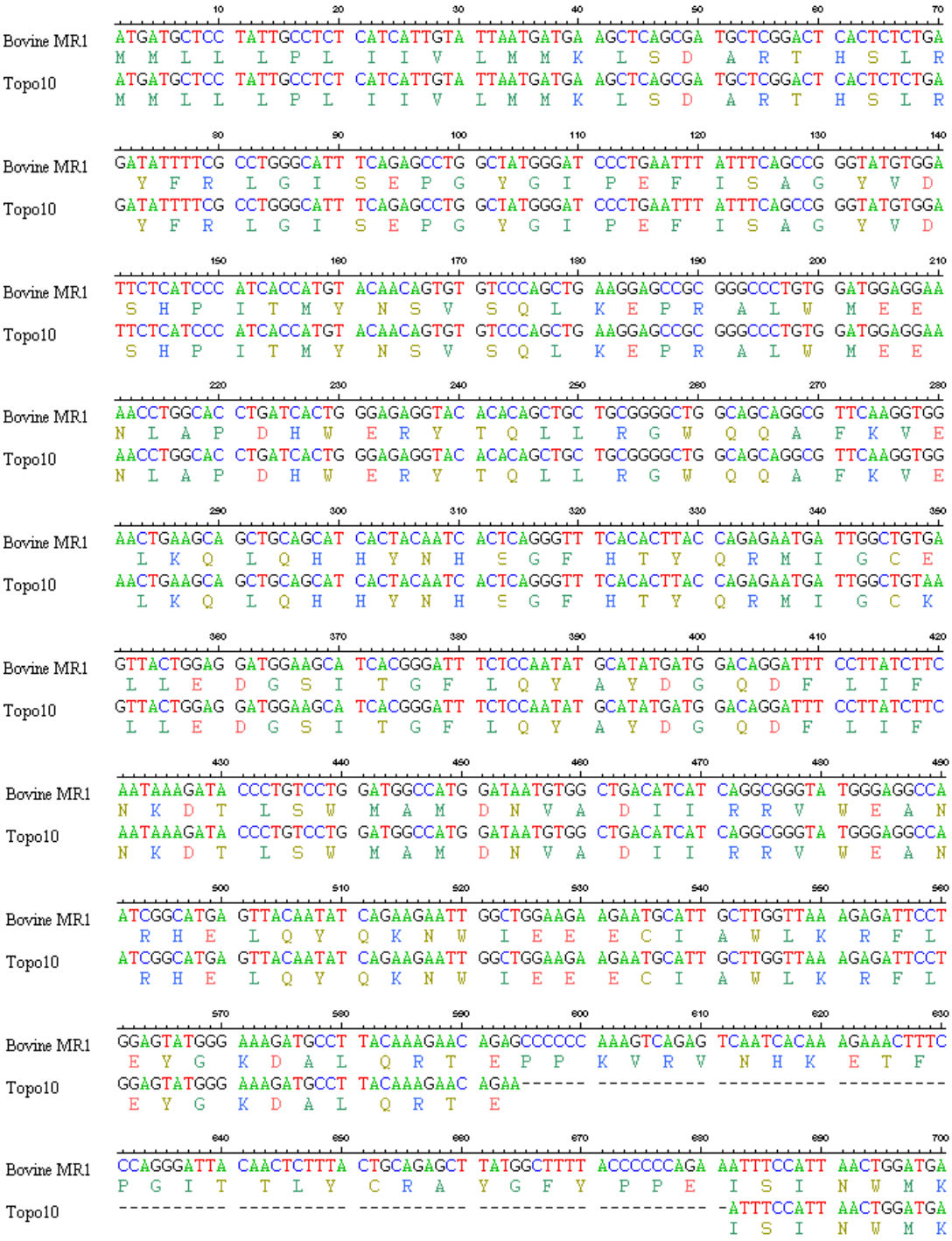


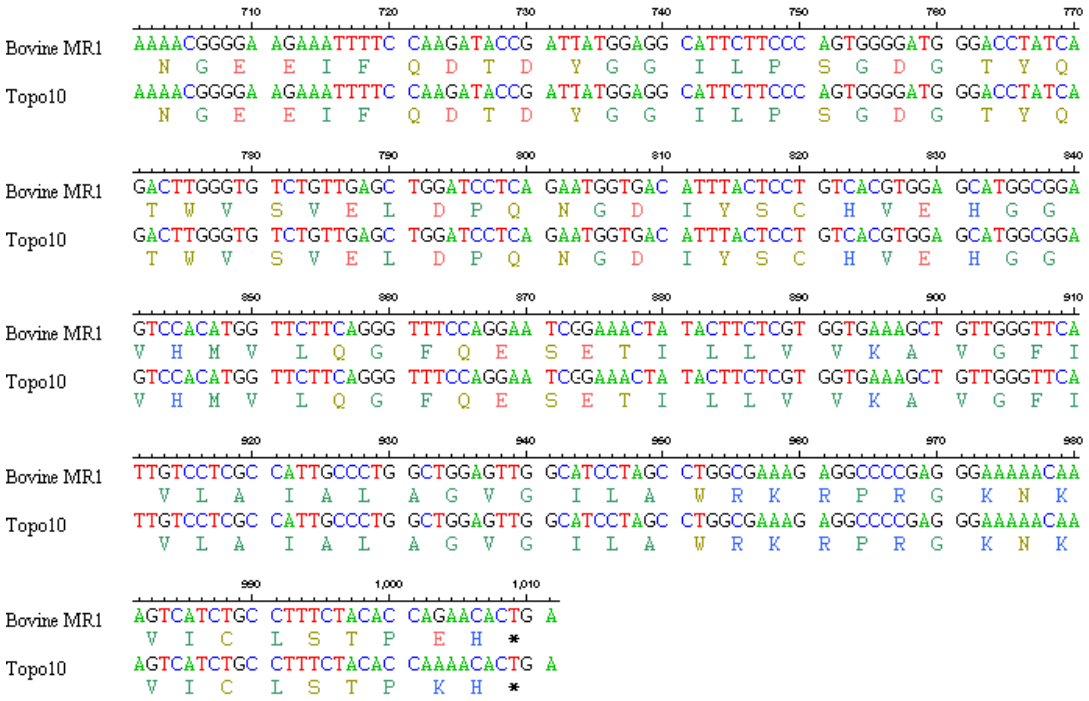

oMR1 110... Sequence
Translate
oMR1 900 ... Sequence
Translate
bMR1-110... Sequence
Translate
bMR1-900 ... Sequence
Translate
bMR1-650 ... Sequence
Translate
oMR1 1100 Sequence
Translate
oMR1 900 Sequence
Translate
bMR1-1100 Sequence
Translate
bMR1-900 Sequence
Translate
bMR1-650 Sequence
Translate


780 790 800 810 820 830 840
oMR1 1100 Sequence GACTTGGGTG TCTGTTGAGC TGGATTCTCA GAACGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGC
Translate T W V S V E L D S Q N G D I Y S C H V E H G G
oMR1 900 Sequence GACTTGGGTG TCTGTTGAGC TGGATTCTCA GAACGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGC
Translate T W V S V E L D S Q N G D I Y S C H V E H G G
bMR1-1100 Sequence GACTTGGGTG TCTGTTGAGC TGGATCCTCA GAATGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGA
Translate T W V S V E L D P Q N G D I Y S C H V E H G G
bMR1-900 Sequence GACTTGGGTG TCTGTTGAGC TGGATCCTCA GAATGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGA
Translate T W V S V E L D P Q N G D I Y S C H V E H G G
bMR1-650 Sequence ---------- ---------- ---------- ---------- ---------- ---------- ----------
Translate
850 860 870 880 890 900 910
oMR1 110... Sequence GTCCACATGG TTCTTCCGGG TTTCCAGGAT TCGGAAACTA TGCTTCTGGT GATGAAAGCT GTTGGGTTCA
Translate V H M V L P G F Q D S E T M L L V M K A V G F I
oMR1 900 ... Sequence GTCCACATGG TTCTTCCGGG TTTCCAGGAT TCGGAAACTA TGCTTCTGGT GATGAAAGCT GTTGGGTTCA
Translate V H M V L P G F Q D S E T M L L V M K A V G F I
bMR1-110... Sequence GTCCACATGG TTCTTCAGGG TTTCCAGGAA TCGGAAACTA TACTTCTCGT GGTGAAAGCT GTTGGGTTCA
Translate V H M V L Q G F Q E S E T I L L V V K A V G F I
bMR1-900 ... Sequence GTCCACATGG TTCTTCAGGG TTTCCAGGAA TCGGAAACTA TACTTCTCGT GGTGAAAGCT GTTGGGTTCA
Translate V H M V L Q G F Q E S E T I L L V V K A V G F I
bMR1-650 ... Sequence ---------- ---------- --GACAGGAA TCGGAAACTA TACTTCTCGT GGTGAAAGCT GTTGGGTTCA
Translate Q E S E T I L L V V K A V G F I
920 930 940 950 960 970 980
oMR1 110... Sequence TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCTTAGC CTGGCGAAGG AAGCCCCGAG GGGAAAACAA
Translate V L A I A L A G V G I L A W R R K P R G E N K
oMR1 900 ... Sequence TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCTTAGC CTGGCGAAGG AAGCCCCGAG GGGAAAACAA
Translate V L A I A L A G V G I L A W R R K P R G E N K
bMR1-110... Sequence TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCCTAGC CTGGCGAAAG AGGCCCCGAG GGAAAAACAA
Translate V L A I A L A G V G I L A W R K R P R G K N K
bMR1-900 ... Sequence TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCCTAGC CTGGCGAAAG AGGCCCCGAG GGAAAAACAA
Translate V L A I A L A G V G I L A W R K R P R G K N K
bMR1-650 ... Sequence TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCCTAGC CTGGCGAAAG AGGCCCCGAG GGAAAAACAA
Translate V L A I A L A G V G I L A W R K R P R G K N K
990 1,000 1,010
oMR1 110... Sequence AGTCATCTAC CTTTCTACAC CAGAACACTG A
Translate V I Y L S T P E H *
oMR1 900 ... Sequence AGTCATCTAC CTTTCTACAC CAGAACACTG A
Translate V I Y L S T P E H *
bMR1-110... Sequence AGTCATCTGC CTTTCTACAC CAGAACACTG A
Translate V I C L S T P E H *
bMR1-900 ... Sequence AGTCATCTGC CTTTCTACAC CAGAACACTG A
Translate V I C L S T P E H *
bMR1-650 ... Sequence AGTCATCTGC CTTTCTACAC CAGAACACTG A
Translate V I C L S T P E H *

14 Appendix G: Potential splice variant of bovine MR1

Nucleotide and amino acid sequences for the coding regions of the potential splice variant of bovine MR1 identified through cloning of bovine MR1 in a TOPO TA expression vector, in comparison to full MR1 coding regions:




710 720 730 740 750 760 770
Bovine MR1 AAAACGGGGA AGAAATTTTC CAAGATACCG ATTATGGAGG CATTCTTCCC AGTGGGGATG GGACCTATCA
N G E E I F Q D T D Y G G I L P S G D G T Y Q
Topo10 AAAACGGGGA AGAAATTTTC CAAGATACCG ATTATGGAGG CATTCTTCCC AGTGGGGATG GGACCTATCA
N G E E I F Q D T D Y G G I L P S G D G T Y Q
780 790 800 810 820 830 840
Bovine MR1 GACTTGGGTG TCTGTTGAGC TGGATCCTCA GAATGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGA
T W V S V E L D P Q N G D I Y S C H V E H G G
Topo10 GACTTGGGTG TCTGTTGAGC TGGATCCTCA GAATGGTGAC ATTTACTCCT GTCACGTGGA GCATGGCGGA
T W V S V E L D P Q N G D I Y S C H V E H G G
850 860 870 880 890 900 910
Bovine MR1 GTCCACATGG TTCTTCAGGG TTTCCAGGAA TCGGAAACTA TACTTCTCGT GGTGAAAGCT GTTGGGTTCA
V H M V L Q G F Q E S E T I L L V V K A V G F I
Topo10 GTCCACATGG TTCTTCAGGG TTTCCAGGAA TCGGAAACTA TACTTCTCGT GGTGAAAGCT GTTGGGTTCA
V H M V L Q G F Q E S E T I L L V V K A V G F I
920 930 940 950 960 970 980
Bovine MR1 TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCCTAGC CTGGCGAAAG AGGCCCCGAG GGAAAAACAA
V L A I A L A G V G I L A W R K R P R G K N K
Topo10 TTGTCCTCGC CATTGCCCTG GCTGGAGTTG GCATCCTAGC CTGGCGAAAG AGGCCCCGAG GGAAAAACAA
V L A I A L A G V G I L A W R K R P R G K N K
990 1,000 1,010
Bovine MR1 AGTCATCTGC CTTTCTACAC CAGAACACTG A
V I C L S T P E H *
Topo10 AGTCATCTGC CTTTCTACAC CAAAACACTG A
V I C L S T P K H *

15 Appendix H: Publications from the thesis

Publications in journals:

Nick Goldfinch, Peter Reinink, et al (2010) "Conservation of mucosal associated invariant T (MAIT) cells and the MR1 restriction element in ruminants, and abundance of MAIT cells in spleen" Vet. Res. (2010) **41**(62).

Announcements in conferences:

Goldfinch, N. G., Connelley, T., Morrison, W. I. (2008) "MAIT cells and MR1" (Proceedings of the The Association for Veterinary Training and Research Work annual conference, Scarborough, 2008)

Goldfinch, N. G., Connelley, T., Morrison, W. I. (2007) "MAIT cells and MR1" (Proceedings of the The Association for Veterinary Training and Research Work annual conference, Scarborough, 2007)

Nucleotide sequences submitted in public databases:

Novel nucleotide sequences submitted in NCBI GenBank were given the accession numbers: FJ423039, FJ423040, FJ423041, FJ423042.

Bibliography

Abi-Rached, L., A. Gilles, et al. (2002). "Evidence of en bloc duplication in vertebrate genomes." Nat Genet **31**(1): 100-5.

Akira, S., K. Okazaki, et al. (1987). "Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining." Science **238**(4830): 1134-8.

Alberts, B., A. Johnson, et al. (2002). Molecular Biology of the Cell; Fourth Edition. New York and London, Garland Science.

Allez, M., J. Brimnes, et al. (2002). "Expansion of CD8+ T cells with regulatory function after interaction with intestinal epithelial cells." Gastroenterology **123**(5): 1516-26.

Allison, J. P. and L. L. Lanier (1987). "Structure, function, and serology of the T-cell antigen receptor complex." Annu Rev Immunol **5**: 503-40.

Arden, B., S. P. Clark, et al. (1995). "Human T-cell receptor variable gene segment families." Immunogenetics **42**(6): 455-500.

Arstila, T. P., A. Casrouge, et al. (1999). "A direct estimate of the human alphabeta T cell receptor diversity." Science **286**(5441): 958-61.

Banchereau, J. and R. M. Steinman (1998). "Dendritic cells and the control of immunity." Nature **392**(6673): 245-52.

Band, M., J. H. Larson, et al. (1998). "A radiation hybrid map of BTA23: identification of a chromosomal rearrangement leading to separation of the cattle MHC class II subregions." Genomics **53**(3): 269-75.

Beck, S. and J. Trowsdale (2000). "The human major histocompatability complex: lessons from the DNA sequence." Annu Rev Genomics Hum Genet **1**: 117-37.

Beissert, S., A. Schwarz, et al. (2006). "Regulatory T cells." J Invest Dermatol **126**(1): 15-24.

Benaroch, P., M. Yilla, et al. (1995). "How MHC class II molecules reach the endocytic pathway." Embo J **14**(1): 37-49.

Bendelac, A. (1995). "Mouse NK1+ T cells." Curr Opin Immunol **7**(3): 367-74.

Bensaid, A. and M. Hadam (1991). "Individual antigens of cattle. Bovine CD4 (BoCD4)." Vet Immunol Immunopathol **27**(1-3): 51-4.

Bensaid, A., A. Kaushal, et al. (1989). "Biochemical characterization of activation-associated bovine class I major histocompatibility complex antigens." Anim Genet **20**(3): 241-55.

Beutler, B. (2004). "Innate immunity: an overview." Mol Immunol **40**(12): 845-59.

Birch, J., G. Codner, et al. (2008). "Genomic location and characterisation of nonclassical MHC class I genes in cattle." Immunogenetics **60**(5): 267-73.

Birch, J., C. De Juan Sanjuan, et al. (2008). "Genomic location and characterisation of MIC genes in cattle." Immunogenetics **60**(8): 477-83.

Birch, J., L. Murphy, et al. (2006). "Generation and maintenance of diversity in the cattle MHC class I region." Immunogenetics **58**(8): 670-9.

Bjorkman, P. J. (1997). "MHC restriction in three dimensions: a view of T cell receptor/ligand interactions." Cell **89**(2): 167-70.

Bjorkman, P. J. and P. Parham (1990). "Structure, function, and diversity of class I major histocompatibility complex molecules." Annu Rev Biochem **59**: 253-88.

Blattman, J. N., D. J. Sourdive, et al. (2000). "Evolution of the T cell repertoire during primary, memory, and recall responses to viral infection." J Immunol **165**(11): 6081-90.

Blumberg, R. S. (1998). "II. One size fits all: nonclassical MHC molecules fulfill multiple roles in epithelial cell function." Am J Physiol Gastrointest Liver Physiol **274**(2): G227-231.

Borg, N. A., L. K. Ely, et al. (2005). "The CDR3 regions of an immunodominant T cell receptor dictate the 'energetic landscape' of peptide-MHC recognition." Nat Immunol **6**(2): 171-80.

Borgulya, P., H. Kishi, et al. (1992). "Exclusion and inclusion of alpha and beta T cell receptor alleles." Cell **69**(3): 529-37.

Bosc, N. and M. P. Lefranc (2003). "The mouse (Mus musculus) T cell receptor alpha (TRA) and delta (TRD) variable genes." Dev Comp Immunol **27**(6-7): 465-97.

Bouvier, M. and D. C. Wiley (1994). "Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules." Science **265**(5170): 398-402.

Boysen, C., M. I. Simon, et al. (1997). "Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with bacterial artificial chromosome clones." Genome Res **7**(4): 330-8.

Boyum, A. (1968). "Isolation of mononuclear cells and granulocytes from human blood. Isolation of monuclear cells by one centrifugation, and of granulocytes by combining centrifugation and sedimentation at 1 g." Scand J Clin Lab Invest Suppl **97**: 77-89.

Braud, V. M., D. S. Allan, et al. (1999). "Functions of nonclassical MHC and non-MHC-encoded class I molecules." Curr Opin Immunol **11**(1): 100-8.

Bryant, P. and H. Ploegh (2004). "Class II MHC peptide loading by the professionals." Curr Opin Immunol **16**(1): 96-102.

Cabaniols, J. P., N. Fazilleau, et al. (2001). "Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase." J Exp Med **194**(9): 1385-90.

Casrouge, A., E. Beaudoing, et al. (2000). "Size estimate of the alpha beta TCR repertoire of naive mouse splenocytes." J Immunol **164**(11): 5782-7.

Cerundolo, V., J. D. Silk, et al. (2009). "Harnessing invariant NKT cells in vaccination strategies." Nat Rev Immunol **9**(1): 28-38.

Chothia, C., D. R. Boswell, et al. (1988). "The outline structure of the T-cell alpha beta receptor." Embo J **7**(12): 3745-55.

Clevers, H., N. D. MacHugh, et al. (1990). "Identification of a bovine surface antigen uniquely expressed on CD4-CD8- T cell receptor gamma/delta+ T lymphocytes." Eur J Immunol **20**(4): 809-17.

Correia-Neves, M., C. Waltzinger, et al. (2001). "The shaping of the T cell repertoire." Immunity **14**(1): 21-32.

Croxford, J. L., S. Miyake, et al. (2006). "Invariant V[alpha]19i T cells regulate autoimmune inflammation." Nat Immunol **7**(9): 987-994.

Cunningham, C. P., W. G. Kimpton, et al. (2001). "Thymic export in aged sheep: a continuous role for the thymus throughout pre- and postnatal life." Eur J Immunol **31**(3): 802-11.

Davis, M. M. and P. J. Bjorkman (1988). "T-cell antigen receptor genes and T-cell recognition." Nature **334**(6181): 395-402.

Davis, W. C., W. C. Brown, et al. (1996). "Analysis of monoclonal antibodies specific for the gamma delta TcR." Vet Immunol Immunopathol **52**(4): 275-83.

Davis, W. C., N. D. MacHugh, et al. (1993). "Identification of a monoclonal antibody reactive with the bovine orthologue of CD3 (BoCD3)." Vet Immunol Immunopathol **39**(1-3): 85-91.

Dellabona, P., E. Padovan, et al. (1994). "An invariant V alpha 24-J alpha Q/V beta 11 T cell receptor is expressed in all individuals by clonally expanded CD4-8- T cells." J Exp Med **180**(3): 1171-6.

Dominguez-Gerpe, L. and M. Rey-Mendez (2003). "Evolution of the thymus size in response to physiological and random events throughout life." Microsc Res Tech **62**(6): 464-76.

Ellis, S. (2004). "The cattle major histocompatibility complex: is it unique?" Vet Immunol Immunopathol **102**(1-2): 1-8.

Ellis, S. A. and K. T. Ballingall (1999). "Cattle MHC: evolution in action?" Immunol Rev **167**: 159-68.

Ellis, S. A., E. C. Holmes, et al. (1999). "Variation in the number of expressed MHC genes in different cattle class I haplotypes." Immunogenetics **50**(5-6): 319-28.

Ellis, S. A., K. A. Staines, et al. (1996). "cDNA sequence of cattle MHC class I genes transcribed in serologically defined haplotypes A18 and A31." Immunogenetics **43**(3): 156-9.

Fearon, D. T. and R. M. Locksley (1996). "The instructive role of innate immunity in the acquired immune response." Science **272**(5258): 50-3.

Fischer Lindahl, K. (1993). "Peptide antigen presentation by non-classical MHC class I molecules." Semin Immunol **5**(2): 117-26.

Flajnik, M. F. and M. Kasahara (2001). "Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system." Immunity **15**(3): 351-62.

Fox, L. M., D. G. Cox, et al. (2009). "Recognition of Lyso-Phospholipids by Human Natural Killer T Lymphocytes." PLoS Biol **7**(10): e1000228.

Fries, R., D. Ewald, et al. (2001). "Assessment of the nucleotide sequence variability in the bovine T-cell receptor alpha delta joining gene region." Anim Biotechnol **12**(1): 29-49.

Garboczi, D. N. and W. E. Biddison (1999). "Shapes of MHC restriction." Immunity **10**(1): 1-7.

Garboczi, D. N., P. Ghosh, et al. (1996). "Structure of the complex between human T-cell receptor, viral peptide and HLA-A2." Nature **384**(6605): 134-41.

Garcia, K. C. and E. J. Adams (2005). "How the T Cell Receptor Sees Antigen--A Structural View." Cell **122**(3): 333-336.

Garcia, K. C., M. Degano, et al. (1998). "Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen." Science **279**(5354): 1166-72.

Garcia, K. C., M. Degano, et al. (1996). "An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex." Science **274**(5285): 209-19.

Garcia, K. C., L. Teyton, et al. (1999). "Structural basis of T cell recognition." Annu Rev Immunol **17**: 369-97.

Gay, C. C. (1965). "Escherichia coli and Neonatal Disease of Calves." Microbiol. Mol. Biol. Rev. **29**(1): 75-101.

Giudicelli, V., D. Chaume, et al. (2005). "IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes." Nucleic Acids Res **33**(Database issue): D256-61.

Giudicelli, V. and M. P. Lefranc (1999). "Ontology for immunogenetics: the IMGT-ONTOLOGY." Bioinformatics **15**(12): 1047-54.

Glusman, G., L. Rowen, et al. (2001). "Comparative genomics of the human and mouse T cell receptor loci." Immunity **15**(3): 337-49.

Goddeeris, B. M. and W. I. Morrison (1988). "Techniques for the generation, cloning, and characterisation of bovine cytotoxic T-cells specific for the protozoan *Theileria parva*." Journal of Tissue Culture Methods **11**: 101-110.

Godfrey, D. I., K. J. Hammond, et al. (2000). "NKT cells: facts, functions and fallacies." Immunol Today **21**(11): 573-83.

Godfrey, D. I., H. R. MacDonald, et al. (2004). "NKT cells: what's in a name?" Nat Rev Immunol **4**(3): 231-7.

Gold, M. C., S. Cerri, et al. (2010). "Human mucosal associated invariant T cells detect bacterially infected cells." PLoS Biol **8**(6): e1000407.

Goldrath, A. W. and M. J. Bevan (1999). "Selecting and maintaining a diverse T-cell repertoire." Nature **402**(6759): 255-62.

Gozalbo-Lopez, B., M. Gomez del Moral, et al. (2009). "The MHC-related protein 1 (MR1) is expressed by a subpopulation of CD38+, IgA+ cells in the human intestinal mucosa." Histol Histopathol **24**(11): 1439-49.

Gromme, M. and J. Neefjes (2002). "Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways." Mol Immunol **39**(3-4): 181-202.

Gumperz, J. E. and M. B. Brenner (2001). "CD1-specific T cells in microbial immunity." Curr Opin Immunol **13**(4): 471-8.

Gumperz, J. E., S. Miyake, et al. (2002). "Functionally distinct subsets of CD1d-restricted natural killer T cells revealed by CD1d tetramer staining." J Exp Med **195**(5): 625-36.

Han, M., L. Harrison, et al. (1999). "Invariant or highly conserved TCR alpha are expressed on double-negative (CD3+CD4-CD8-) and CD8+ T cells." J Immunol **163**(1): 301-11.

Hansen, T. H., S. Huang, et al. (2007). "Patterns of nonclassical MHC antigen presentation." Nat Immunol **8**(6): 563-568.

Harty, J. T., A. R. Tvinnereim, et al. (2000). "CD8+ T cell effector mechanisms in resistance to infection." Annu Rev Immunol **18**: 275-308.

Hashimoto, K., M. Hirai, et al. (1995). "A Gene Outside the Human MHC Related to Classical HLA Class I Genes." Science **269**(5224): 693-695.

Haynes, M. R. and G. E. Wu (2004). "Evolution of the variable gene segments and recombination signal sequences of the human T-cell receptor alpha/delta locus." Immunogenetics **56**(7): 470-9.

Haynes, M. R. and G. E. Wu (2007). "Gene discovery at the human T-cell receptor alpha/delta locus." Immunogenetics **59**(2): 109-21.

Heath, W. R. and J. F. Miller (1993). "Expression of two alpha chains on the surface of T cells in T cell receptor transgenic mice." J Exp Med **178**(5): 1807-11.

Hein, W. R. and C. R. Mackay (1991). "Prominence of gamma delta T cells in the ruminant immune system." Immunol Today **12**(1): 30-4.

Hein, W. R., A. Marcuz, et al. (1991). "Primary structure of the sheep T-cell receptor alpha chain." Immunogenetics **34**(1): 39-41.

Hennecke, J. and D. C. Wiley (2001). "T cell receptor-MHC interactions up close." Cell **104**(1): 1-4.

Hosono, M., E. Toichi, et al. (1995). "Development of autoreactivity and changes of T cell repertoire in different strains of aging mice." Mech Ageing Dev **78**(3): 197-214.

Huang, S., S. Gilfillan, et al. (2005). "Evidence for MR1 Antigen Presentation to Mucosal-associated Invariant T Cells." J. Biol. Chem. **280**(22): 21183-21193.

Huang, S., S. Gilfillan, et al. (2008). "MR1 uses an endocytic pathway to activate mucosal-associated invariant T cells." J. Exp. Med. **205**(5): 1201-1211.

Huang, S., E. Martin, et al. (2009). "MR1 antigen presentation to mucosal-associated invariant T cells was highly conserved in evolution." Proceedings of the National Academy of Sciences: -.

Huseby, E. S., J. White, et al. (2005). "How the T cell repertoire becomes peptide and MHC specific." Cell **122**(2): 247-60.

Illes, Z., T. Kondo, et al. (2000). "Differential expression of NK T cell V alpha 24J alpha Q invariant TCR chain in the lesions of multiple sclerosis and chronic inflammatory demyelinating polyneuropathy." J Immunol **164**(8): 4375-81.

Illes, Z., M. Shimamura, et al. (2004). "Accumulation of V{alpha}7.2-J{alpha}33 invariant T cells in human autoimmune inflammatory lesions in the nervous system." Int. Immunol. **16**(2): 223-230.

Ishiguro, N., A. Tanaka, et al. (1990). "Sequence analysis of bovine T-cell receptor alpha chain." Immunogenetics **31**(1): 57-60.

Ishihara, S., M. Nieda, et al. (1999). "CD8(+)NKR-P1A (+)T cells preferentially accumulate in human liver." Eur J Immunol **29**(8): 2406-13.

Janeway, C. (2001). Immunobiology; Fifth Edition. New York and London, Garland Science.

Janeway, C. A., Jr. (1993). "How the immune system recognizes invaders." Sci Am **269**(3): 72-9.

Janeway, C. A., Jr. and R. Medzhitov (2002). "Innate immune recognition." Annu Rev Immunol **20**: 197-216.

Kampen, A. H., I. Olsen, et al. (2006). "Lymphocyte subpopulations and neutrophil function in calves during the first 6 months of life." Vet Immunol Immunopathol **113**(1-2): 53-63.

Kasahara, M. (1999). "Genome dynamics of the major histocompatibility complex: insights from genome paralogy." Immunogenetics **50**(3-4): 134-45.

Kasahara, M., M. Hayashi, et al. (1996). "Chromosomal localization of the proteasome Z subunit gene reveals an ancient chromosomal duplication involving the major histocompatibility complex." Proc Natl Acad Sci U S A **93**(17): 9096-101.

Kawachi, I., J. Maldonado, et al. (2006). "MR1-Restricted V{alpha}19i Mucosal-Associated Invariant T Cells Are Innate T Cells in the Gut Lamina Propria That Provide a Rapid and Diverse Cytokine Response." J Immunol **176**(3): 1618-1627.

Kawano, T., J. Cui, et al. (1997). "CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides." Science **278**(5343): 1626-9.

Kazatchkine, M. D. (1995). "Nomenclature for T-cell receptor (TCR) gene segments of the immune system. WHO-IUIS Nomenclature Sub-Committee on TCR Designation." Immunogenetics **42**(6): 451-3.

Kedl, R. M., B. C. Schaefer, et al. (2002). "T cells down-modulate peptide-MHC complexes on APCs in vivo." Nat Immunol **3**(1): 27-32.

Kelley, J., L. Walter, et al. (2005). "Comparative genomics of major histocompatibility complexes." Immunogenetics **56**(10): 683-95.

Kidd, P. (2003). "Th1/Th2 balance: the hypothesis, its limitations, and implications for health and disease." Altern Med Rev **8**(3): 223-46.

Kim, S., S. Lalani, et al. (2008). "Impact of bacteria on the phenotype, functions, and therapeutic activities of invariant NKT cells in mice." J Clin Invest **118**(6): 2301-15.

Koch, J. and R. Tampe (2006). "The macromolecular peptide-loading complex in MHC class I-dependent antigen presentation." Cell Mol Life Sci **63**(6): 653-62.

Krangel, M. S., M. T. McMurry, et al. (2000). "Accessibility control of T cell receptor gene rearrangement in developing thymocytes. The TCR alpha/delta locus." Immunol Res **22**(2-3): 127-35.

Lantz, O. and A. Bendelac (1994). "An invariant T cell receptor alpha chain is used by a unique subset of major histocompatibility complex class I-specific CD4+ and CD4-8- T cells in mice and humans." J Exp Med **180**(3): 1097-106.

Le Bourhis, L., E. Martin, et al. (2010). "Antimicrobial activity of mucosal-associated invariant T cells." Nat Immunol **11**(8): 701-8.

Lefranc, M. P., V. Giudicelli, et al. (1998). "IMGT, the International ImMunoGeneTics database." Nucleic Acids Res **26**(1): 297-303.

Levelt, C. N., B. Wang, et al. (1995). "Regulation of T cell receptor (TCR)-beta locus allelic exclusion and initiation of TCR-alpha locus rearrangement in immature thymocytes by signaling through the CD3 complex." Eur J Immunol **25**(5): 1257-61.

Lin, H., M. Nieda, et al. (2006). "Comparative gene expression analysis of NKT cell subpopulations." J Leukoc Biol **80**(1): 164-73.

Linsen, L., V. Somers, et al. (2005). "Immunoregulation of autoimmunity by natural killer T cells." Hum Immunol **66**(12): 1193-202.

Livak, F. (2004). "In vitro and in vivo studies on the generation of the primary T-cell receptor repertoire." Immunol Rev **200**: 23-35.

Livak, F. and D. G. Schatz (1996). "T-cell receptor alpha locus V(D)J recombination by-products are abundant in thymocytes and mature T cells." Mol Cell Biol **16**(2): 609-18.

Looringh van Beeck, F. A., P. Reinink, et al. (2009). "Functional CD1d and/or NKT cell invariant chain transcript in horse, pig, African elephant and guinea pig, but not in ruminants." Mol Immunol **46**(7): 1424-31.

Lustig, A., A. Carter, et al. (2009). "Transcriptome analysis of murine thymocytes reveals age-associated changes in thymic gene expression." Int J Med Sci **6**(1): 51-64.

MacHugh, N. D. and P. Sopp (1991). "Individual antigens of cattle. Bovine CD8 (BoCD8)." Vet Immunol Immunopathol **27**(1-3): 65-9.

Mackelprang, R., R. J. Livingston, et al. (2006). "Sequence diversity, natural selection and linkage disequilibrium in the human T cell receptor alpha/delta locus." Hum Genet **119**(3): 255-66.

Madden, D. R. (1995). "The three-dimensional structure of peptide-MHC complexes." Annu Rev Immunol **13**: 587-622.

Maenaka, K. and E. Y. Jones (1999). "MHC superfamily structure and the immune system." Curr Opin Struct Biol **9**(6): 745-53.

Makino, Y., R. Kanno, et al. (1995). "Predominant expression of invariant V alpha 14+ TCR alpha chain in NK1.1+ T cell populations." Int Immunol **7**(7): 1157-61.

Manning, T. C., C. J. Schlueter, et al. (1998). "Alanine scanning mutagenesis of an alphabeta T cell receptor: mapping the energy of antigen recognition." Immunity **8**(4): 413-25.

Martin, E., E. Treiner, et al. (2009). "Stepwise Development of MAIT Cells in Mouse and Human." PLoS Biol **7**(3): e1000054.

Mathis, D. and C. Benoist (2004). "Back to central tolerance." Immunity **20**(5): 509-16.

Mayer, G. (2007). Microbiology and Immunology On-line. Immunology, http://pathmicro.med.sc.edu/book/immunol-sta.htm R. C. Hunt, University of South Carolina School of Medicine.

Maziarz, R. T., J. Fraser, et al. (1986). "The human HLA-specific monoclonal antibody W6/32 recognizes a discontinuous epitope within the alpha 2 domain of murine H-2Db." Immunogenetics **24**(3): 206-8.

McBlane, J. F., D. C. van Gent, et al. (1995). "Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps." Cell **83**(3): 387-95.

McHeyzer-Williams, M. G. and M. M. Davis (1995). "Antigen-specific development of primary and memory T cells in vivo." Science **268**(5207): 106-11.

Medzhitov, R. and C. Janeway, Jr. (2000). "Innate immunity." N Engl J Med **343**(5): 338-44.

Miley, M. J., S. M. Truscott, et al. (2003). "Biochemical Features of the MHC-Related Protein 1 Consistent with an Immunological Function." J Immunol **170**(12): 6090-6098.

Montoya, C. J., D. Pollard, et al. (2007). "Characterization of human invariant natural killer T subsets in health and disease using a novel invariant natural

killer T cell-clonotypic monoclonal antibody, 6B11." Immunology **122**(1): 1-14.

Morrison, W. I. and W. C. Davis (1991). "Individual antigens of cattle. Differentiation antigens expressed predominantly on CD4- CD8- T lymphocytes (WC1, WC2)." Vet Immunol Immunopathol **27**(1-3): 71-6.

Moseley, W., M. Watson, et al. (1989). "CD1 defines conserved linkage group border between human chromosomes 1 and mouse chromosomes 1 and 3." Immunogenetics **30**(5): 378-382.

Moss, P. A. and J. I. Bell (1995). "Sequence analysis of the human alpha beta T-cell receptor CDR3 region." Immunogenetics **42**(1): 10-8.

Murray, J. S. (1998). "How the MHC selects Th1/Th2 immunity." Immunology Today **19**(4): 157-162.

Neveu, B., K. Echasserieau, et al. (2006). "Impact of CD8-MHC class I interaction in detection and sorting efficiencies of antigen-specific T cells using MHC class I/peptide multimers: contribution of pMHC valency." Int Immunol **18**(7): 1139-45.

Nieuwenhuis, P. and D. Opstelten (1984). "Functional anatomy of germinal centers." Am J Anat **170**(3): 421-35.

Nikolich-Zugich, J., M. K. Slifka, et al. (2004). "The many important facets of T-cell repertoire diversity." Nat Rev Immunol **4**(2): 123-32.

Okamoto, N., O. Kanie, et al. (2005). "Synthetic alpha-mannosyl ceramide as a potent stimulant for an NKT cell repertoire bearing the invariant Valpha19-Jalpha26 TCR alpha chain." Chem Biol **12**(6): 677-83.

Padovan, E., G. Casorati, et al. (1993). "Expression of two T cell receptor alpha chains: dual receptor T cells." Science **262**(5132): 422-4.

Parra-Cuadrado, J. F., M. G. del Moral, et al. (2001). "Characterization of the MHC class I-related MR1 locus in nonhuman primates." Immunogenetics **53**(8): 643-8.

Parra-Cuadrado, J. F., P. Navarro, et al. (2000). "A study on the polymorphism of human MHC class I-related MR1 gene and identification of an MR1-like pseudogene." Tissue Antigens **56**(2): 170-2.

Pellicci, D. G., O. Patel, et al. (2009). "Differential recognition of CD1d-alpha-galactosyl ceramide by the V beta 8.2 and V beta 7 semi-invariant NKT T cell receptors." Immunity **31**(1): 47-59.

Peterfalvi, A., E. Gomori, et al. (2008). "Invariant V{alpha}7.2-J{alpha}33 TCR is expressed in human kidney and brain tumors indicating infiltration by mucosal-associated invariant T (MAIT) cells." Int. Immunol.: dxn111.

Pietra, G., C. Romagnani, et al. (2009). "HLA-E and HLA-E-bound peptides: recognition by subsets of NK and T cells." Curr Pharm Des **15**(28): 3336-44.

Porcelli, S., C. E. Yockey, et al. (1993). "Analysis of T cell antigen receptor (TCR) expression by human peripheral blood CD4-8- alpha/beta T cells demonstrates preferential use of several V beta genes and an invariant TCR alpha chain." J Exp Med **178**(1): 1-16.

Probst, J., S. G. Blumenthal, et al. (2004). "A conserved sequence in the mouse variable T cell receptor alpha recombination signal sequence 23-bp spacer can affect recombination." Eur J Immunol **34**(8): 2179-90.

Raulet, D. H. (1989). "The structure, function, and molecular genetics of the gamma/delta T cell receptor." Annu Rev Immunol **7**: 175-207.

Reinink, P. and I. Van Rhijn (2009). "The bovine T cell receptor alpha/delta locus contains over 400 V genes and encodes V genes without CDR2." Immunogenetics **(2009) 61:541–549**: 541-549.

Reis e Sousa, C. (2004). "Activation of dendritic cells: translating innate into adaptive immunity." Curr Opin Immunol **16**(1): 21-5.

Rhind, S. M., J. Hopkins, et al. (1999). "Amino-terminal sequencing of sheep CD1 antigens and identification of a sheep CD1D gene." Immunogenetics **49**(3): 225-30.

Riegert, P., V. Wanner, et al. (1998). "Genomics, Isoforms, Expression, and Phylogeny of the MHC Class I-Related MR1 Gene." J Immunol **161**(8): 4066-4077.

Rock, K. L. (1996). "A new foreign policy: MHC class I molecules monitor the outside world." Immunol Today **17**(3): 131-7.

Rock, K. L., I. A. York, et al. (2004). "Post-proteasomal antigen processing for major histocompatibility complex class I presentation." Nat Immunol **5**(7): 670-7.

Rock, K. L., I. A. York, et al. (2002). "Protein degradation and the generation of MHC class I-presented peptides." Adv Immunol **80**: 1-70.

Roy, K. C., I. Maricic, et al. (2008). "Involvement of Secretory and Endosomal Compartments in Presentation of an Exogenous Self-Glycolipid to Type II NKT Cells." J Immunol **180**(5): 2942-2950.

Rudolph, M. G., R. L. Stanfield, et al. (2006). "How TCRs bind MHCs, peptides, and coreceptors." Annu Rev Immunol **24**: 419-66.

Satyanarayana, K., S. Hata, et al. (1988). "Genomic organization of the human T-cell antigen-receptor alpha/delta locus." Proc Natl Acad Sci U S A **85**(21): 8166-70.

Savage, P. A., J. J. Boniface, et al. (1999). "A kinetic basis for T cell receptor repertoire selection during an immune response." Immunity **10**(4): 485-92.

Scaviner, D. and M. P. Lefranc (2000). "The human T cell receptor alpha variable (TRAV) genes." Exp Clin Immunogenet **17**(2): 83-96.

Schatz, D. G. (2004). "V(D)J recombination." Immunol Rev **200**: 5-11.

Schlissel, M. S. (2003). "Regulating antigen-receptor gene assembly." Nat Rev Immunol **3**(11): 890-9.

Schmutz, S. M., J. M. Stookey, et al. (2001). "A QTL Study of Cattle Behavioral Traits in Embryo Transfer Families." J Hered **92**(3): 290-292.

Schrenzel, M. D. and D. A. Ferrick (1995). "Horse (Equus caballus) T-cell receptor alpha, gamma, and delta chain genes: nucleotide sequences and tissue-specific gene expression." Immunogenetics **42**(2): 112-22.

Schumann, J. and G. De Libero (2007). "MR1-restricted Valpha19i T cells: a second population recognizing lipid antigens?" Eur J Immunol **37**(7): 1724-6.

Sebzda, E., S. Mariathasan, et al. (1999). "Selection of the T cell repertoire." Annu Rev Immunol **17**: 829-74.

Seino, K.-i. and M. Taniguchi (2005). "Functionally distinct NKT cell subsets and subtypes." J. Exp. Med. **202**(12): 1623-1626.

Shiina, T., A. Ando, et al. (2001). "Genomic anatomy of a premier major histocompatibility complex paralogous region on chromosome 1q21-q22." Genome Res **11**(5): 789-802.

Shimamura, M., Y.-Y. Huang, et al. (2008). "Altered production of immunoregulatory cytokines by invariant V{alpha}19 TCR-bearing cells dependent on the duration and intensity of TCR engagement." Int. Immunol.: dxn136.

Shimamura, M., Y.-Y. Huang, et al. (2007). "Modulation of Va19 NKT cell immune responses by a-mannosyl ceramide derivatives consisting of a series of modified sphingosines." European Journal of Immunology **37**(7): 1836-1844.

Smith, W. D., F. Jackson, et al. (1981). "Changes in the flow and composition of gastric lymph in sheep repeatedly infected with Ostertagia circumcincta." J Comp Pathol **91**(4): 553-64.

Starr, T. K., S. C. Jameson, et al. (2003). "Positive and negative selection of T cells." Annu Rev Immunol **21**: 139-76.

Steffens, C. M., L. Al-Harthi, et al. (2000). "Evaluation of thymopoiesis using T cell receptor excision circles (TRECs): differential correlation between adult and pediatric TRECs and naive phenotypes." Clin Immunol **97**(2): 95-101.

Steinle, A., P. Li, et al. (2001). "Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family." Immunogenetics **53**(4): 279-87.

Stephens, H. A. (2001). "MICA and MICB genes: can the enigma of their polymorphism be resolved?" Trends Immunol **22**(7): 378-85.

Sun, J., D. J. Leahy, et al. (1995). "Interaction between CD8 and major histocompatibility complex (MHC) class I mediated by multiple contact surfaces that include the alpha 2 and alpha 3 domains of MHC class I." J Exp Med **182**(5): 1275-80.

Sunderkotter, C., H. Kalden, et al. (1997). "Aging and the skin immune system." Arch Dermatol **133**(10): 1256-62.

Takada, K. and S. C. Jameson (2009). "Naive T cell homeostasis: from awareness of space to a sense of place." Nat Rev Immunol **9**(12): 823-832.

Takeshima, S.-N. and Y. Aida (2006). "Structure, function and disease susceptibility of the bovine major histocompatibility complex." Animal Science Journal **77**(2): 138-150.

Takihara, Y., D. Tkachuk, et al. (1988). "Sequence and organization of the diversity, joining, and constant region genes of the human T-cell delta-chain locus." Proc Natl Acad Sci U S A **85**(16): 6097-101.

Teng, M. K., A. Smolyar, et al. (1998). "Identification of a common docking topology with substantial variation among different TCR-peptide-MHC complexes." Curr Biol **8**(7): 409-12.

The Bovine Genome Sequencing and Analysis Consortium, C. G. Elsik, et al. (2009). "The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution." Science **324**(5926): 522-528.

The M. H. C. sequencing consortium (1999). "Complete sequence and gene map of a human major histocompatibility complex." Nature **401**(6756): 921-923.

Thompson, S. D., M. Larche, et al. (1992). "Diversity of T-cell receptor alpha gene transcripts in the newborn and adult periphery." Immunogenetics **36**(2): 95-103.

Tilloy, F., E. Treiner, et al. (1999). "An Invariant T Cell Receptor alpha Chain Defines a Novel TAP-independent Major Histocompatibility Complex Class Ib-restricted alpha /beta T Cell Subpopulation in Mammals." J. Exp. Med. **189**(12): 1907-1921.

Toye, P. G., N. D. MacHugh, et al. (1990). "Transfection into mouse L cells of genes encoding two serologically and functionally distinct bovine class I MHC molecules from a MHC-homozygous animal: evidence for a second class I locus in cattle." Immunology **70**(1): 20-6.

Treiner, E., L. Duban, et al. (2003). "Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1." Nature **422**(6928): 164-169.

Treiner, E., L. Duban, et al. (2005). "Mucosal-associated invariant T (MAIT) cells: an evolutionarily conserved T cell subset." Microbes and Infection **7**(3): 552-559.

Treiner, E. and O. Lantz (2006). "CD1d- and MR1-restricted invariant T cells: of mice and men." Current Opinion in Immunology **18**(5): 519-526.

Turner, S. J., P. C. Doherty, et al. (2006). "Structural determinants of T-cell receptor bias in immunity." Nat Rev Immunol **6**(12): 883-94.

Uematsu, Y., S. Ryser, et al. (1988). "In transgenic mice the introduced functional T cell receptor beta gene prevents expression of endogenous beta genes." Cell **52**(6): 831-41.

Uenishi, H., H. Hiraiwa, et al. (2003). "Genomic structure around joining segments and constant regions of swine T-cell receptor alpha/delta (TRA/TRD) locus." Immunology **109**(4): 515-26.

van Ewijk, W., J. Rozing, et al. (1977). "Cellular events during the primary immune response in the spleen. A fluorescence- light- and electronmicroscopic study in germfree mice." Cell Tissue Res **183**(4): 471-89.

Van Rhijn, I., A. P. Koets, et al. (2006). "The bovine CD1 family contains group 1 CD1 proteins, but no functional CD1d." J Immunol **176**(8): 4888-93.

Van Rhijn, I., R. Spiering, et al. (2007). "Highly diverse TCR delta chain repertoire in bovine tissues due to the use of up to four D segments per delta chain." Mol Immunol **44**(12): 3155-61.

Walter, L. and E. Gunther (1998). "Isolation and molecular characterization of the rat MR1 homologue, a non-MHC-linked class I-related gene." Immunogenetics **47**(6): 477-82.

Walzer, T., M. Dalod, et al. (2005). "Natural-killer cells and dendritic cells: "l'union fait la force"." Blood **106**(7): 2252-2258.

Wheeler, C. J., P. von Hoegen, et al. (1992). "An immunological role for the CD8 beta-chain." Nature **357**(6375): 247-9.

Womack, J. E. and S. R. Kata (1995). "Bovine genome mapping: evolutionary inference and the power of comparative genomics." Curr Opin Genet Dev **5**(6): 725-33.

Womack, J. E. and Y. D. Moll (1986). "Gene map of the cow: conservation of linkage with mouse and man." J Hered **77**(1): 2-7.

Wu, L. C., D. S. Tuot, et al. (2002). "Two-step binding mechanism for T-cell receptor recognition of peptide MHC." Nature **418**(6897): 552-6.

Xie, T., L. Rowen, et al. (2003). "Analysis of the gene-dense major histocompatibility complex class III region and its comparison to mouse." Genome Res **13**(12): 2621-36.

Xu, Y., L. Davidson, et al. (1996). "Function of the pre-T-cell receptor alpha chain in T-cell development and allelic exclusion at the T-cell receptor beta locus." Proc Natl Acad Sci U S A **93**(5): 2169-73.

Yamaguchi, H. and K. Hashimoto (2002). "Association of MR1 Protein, an MHC Class I-Related Molecule, with [beta]2-Microglobulin." Biochemical and Biophysical Research Communications **290**(2): 722-729.

Yamaguchi, H., M. Hirai, et al. (1997). "A highly conserved major histocompatibility complex class I-related gene in mammals." Biochem Biophys Res Commun **238**(3): 697-702.

Yamaguchi, H., Y. Kurosawa, et al. (1998). "Expanded Genomic Organization of Conserved Mammalian MHC Class I-Related Genes, Human MR1 and Its Murine Ortholog." Biochemical and Biophysical Research Communications **250**: 558-564.

Yewdell, J. W., E. Reits, et al. (2003). "Making sense of mass destruction: quantitating MHC class I antigen presentation." Nat Rev Immunol **3**(12): 952-61.

Yung Yu, C., Z. Yang, et al. (2000). "The human and mouse MHC class III region: a parade of 21 genes at the centromeric segment." Immunology Today **21**(7): 320-328.

Zhou, D., J. Mattner, et al. (2004). "Lysosomal glycosphingolipid recognition by NKT cells." Science **306**(5702): 1786-9.

Zinkernagel, R. M. and P. C. Doherty (1974). "Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system." Nature **248**(450): 701-2.